GENOMIC PREDICTION AND GENOME WIDE ASSOCIATION MAPPING FOR DISEASE RESISTANCE IN WHEAT

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Philomin Juliana

January 2017

GENOMIC PREDICTION AND GENOME WIDE ASSOCIATION MAPPING FOR DISEASE RESISTANCE IN WHEAT

Philomin Juliana, Ph.D.

Cornell University 2017

Wheat (*Triticum aestivum L.*) is one of the major food crops in the world that is grown on more land area than any other commercial crop. The demand for wheat is expected to increase by 60% by 2050 which cannot be met with the current yield gain of 1%. Hence, it is important to evaluate different strategies for increasing the genetic gain in wheat. With this focus, we evaluated two strategies, genomic prediction and genome-wide association studies (GWAS) for disease resistance in CIMMYT's international bread wheat screening nurseries (IBWSN). Our objective was to compare different prediction models for resistance to leaf rust (LR), stem rust (SR), stripe rust (STR), Septoria tritici blotch (STB), Stagonospora nodorum blotch (SNB) and tan spot (TS) in the 45th and 46th IBWSN entries. The prediction models tested include: Least-squares (LS), genomic-BLUP (G-BLUP), Bayesian ridge regression (BRR), Bayes A (BA), Bayes B (BB), Bayes Cπ (BC), Bayesian least absolute shrinkage and selection operator (BL), reproducing kernel Hilbert spaces (RKHS) markers (RKHS-M), RKHS pedigree (RKHS-P) and RKHS markers and pedigree (RKHS-MP). The 333 lines in the 45th IBWSN and the 313 lines in the 46th IBWSN were genotyped using genotyping-by-sequencing markers. For the rusts, the mean prediction accuracies were 0.74 for LR seedling, 0.56 for LR APR, 0.65 for SR APR, 0.78 for YR seedling and 0.71 for YR APR. For the leaf spotting diseases, the mean genomic prediction accuracies were 0.45 for STB APR, 0.55 for SNB seedling, 0.66 for TS seedling and 0.48 for TS APR. Using genome-wide marker based models resulted in an average of 42-48% increase in accuracy over LS. Overall, the RKHS-MP model gave the highest accuracies, while LS gave the lowest. GWAS was also

performed on these traits and several significant markers and candidate genes were identified. We conclude that implementing GWAS and genomic selection in breeding for these diseases would help to achieve higher accuracies and rapid gains from selection.

BIOGRAPHICAL SKETCH

"The LORD lifts the poor from the dust and the needy from the garbage dump. HE sets them among princes, placing them in seats of honor" - 1 Samuel 2:8.

Philomin's life and educational career are a testimony to the goodness, greatness and faithfulness of the LORD JESUS, WHO lifted her up from the dust to unimaginable places. She was born on 23rd September, 1989 in Tuticorin, Tamil Nadu, India to her parents, Mr. Johnson and Mrs. Daisy Johnson. Right from her childhood, the LORD lead her and blessed her in all that she did. Her schooling in Holy Cross Anglo-Indian higher secondary school gave her a strong academic foundation and she was motivated by Proverbs 4:13, "Always remember what you have learned. Your education is your life-guard it well". Her fascination for plant sciences made her pursue a Bachelors in plant biotechnology from Tamil Nadu Agricultural University, Coimbatore, India from 2007-2011. She then pursued a Masters in Plant Breeding and Genetics at Cornell University funded by the Cornell-Sathguru foundation for development and Masters in Biotechnology Business Management from Tamil Nadu Agricultural University, Coimbatore. Her love for wheat breeding and genetics effloresced during her Masters research that involved mapping genes for traits associated with durable rust resistance under the outstanding mentorship of Dr. Mark Sorrells and Dr. Ronnie Coffman. The incredibly magnificent wheat genome and the intriguing mechanism of durable disease resistance enthralled her and transformed her interest in research into a passion. She then received Monsanto's Beachell and Borlaug International scholars fellowship to pursue a Ph.D. in Plant Breeding and Genetics at Cornell University under the excellent guidance of Dr. Mark Sorrells, Dr. Gary Bergstrom and Dr. Ronnie Coffman. She also had the wonderful opportunity to collaborate with Dr. Ravi Singh, Dr. Pawan Singh and Dr. Jose Crossa from CIMMYT. During her Ph.D., she has visited CIMMYT several times and was also part of the wheat improvement training in March, 2016. While a doctoral student at Cornell, Philomin Juliana prepared this dissertation as partial fulfillment of the Doctor of Philosophy degree. It is expected that her degree will be conferred by the College of Agriculture and Life Sciences in January, 2017.

Dedicated to my LORD and SAVIOR JESUS CHRIST

ACKNOWLEDGMENTS

I am greatly indebted to my LORD JESUS CHRIST who sustained me and carried me in HIS loving arms throughout my doctoral studies. HE was my fountain of strength, my hope, my joy, my Prince of peace and my everything throughout my Ph.D. Without HIM, I would not have been here and this degree would have not been possible.

I would also like to express my sincere gratitude to my advisor, Dr. Mark Sorrells, for his outstanding mentorship. He has always been such a great inspiration and motivated me in all my endeavors. I also wish to extend my heartfelt gratitude to my committee members, Dr. Gary Bergstrom and Dr. Ronnie Coffman for their wonderful encouragement and support throughout my Ph.D. I consider it such an honor and privilege to be mentored by these three excellent advisors, who were incredibly kind and helpful to me, always. I am also very grateful to Dr. Ravi Singh, Dr. Pawan Singh and Dr. Jose Crossa, who provided immense support for my research at CIMMYT. A special thanks to Dr. Jean-Luc Jannink and Dr. Jessica Rutkoski, who provided very helpful comments. I would also like to express my gratitude to all my lab members and the plant breeding faculty who were always very supportive.

I am also very grateful to Monsanto's Beachell and Borlaug International Scholars program for funding my Ph.D. and research. It was instrumental in shaping my research and professional career and I am very thankful to Dr. Edward Runge for coordinating it. I am also very grateful to my parents who always supported me and cared for me.

TABLE OF CONTENTS

CHAPTER 3

GENOME WIDE ASSOCIATION STUDIES FOR RESISTANCE TO LEAF RUST, STRIPE RUST, STAGONOSPORA NODORUM BLOTCH AND TAN SPOT IN WHEAT REVEALS POTENTIAL CANDIDATE GENES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

GENOMIC AND PEDIGREE BASED PREDICTION FOR LEAF, STEM AND STRIPE RUST RESISTANCE IN WHEAT

ABSTRACT

The unceasing plant-pathogen arms race and ephemeral nature of some rust resistance genes have been challenging for wheat (*Triticum aestivum* L.) breeding programs and farmers. Hence, it is important to devise strategies for effective evaluation and exploitation of quantitative rust resistance. One promising approach that could accelerate gain from selection for rust resistance is 'genomic selection' which utilizes dense genome-wide markers to estimate the breeding values (BVs) for quantitative traits. Our objective was to compare three genomic prediction models including genomic best linear unbiased prediction (GBLUP), GBLUP-A that was GBLUP with selected loci as fixed effects and reproducing kernel Hilbert spaces-markers (RKHS-M) with least squares (LS) approach, RKHS-pedigree (RKHS-P) and RKHS-markers and pedigree (RKHS-MP) to determine the BVs for seedling and/or adult plant resistance (APR) to leaf rust (LR), stem rust (SR) and stripe rust (YR). The 333 lines in the 45th IBWSN and the 313 lines in the 46th IBWSN were genotyped using genotyping-by-sequencing and phenotyped in replicated trials. The mean prediction accuracies ranged from 0.31-0.74 for LR seedling, 0.12-0.56 for LR APR, 0.31-0.65 for SR APR, 0.70-0.78 for YR seedling and 0.34-0.71 for YR APR. For most datasets, the RKHS-MP model gave the highest accuracies, while LS gave the lowest. GBLUP, GBLUP A, RKHS-M and RKHS-P models gave similar accuracies. Using genome-wide marker based models resulted in an average of 42% increase in accuracy over LS. We conclude that GS is a promising approach for improvement of quantitative rust resistance and can be

implemented in the breeding pipeline.

ABBREVIATIONS

APR, adult plant resistance; BLUP, best linear unbiased prediction; BVs, breeding values; CIMMYT, Centro Internacional de Mejoramiento de Maíz y Trigo; GBLUP, genomic best linear unbiased prediction; GBLUP A, genomic best linear unbiased prediction with selected loci as fixed effects; GBS, genotyping-by-sequencing; HWWAMP, hard winter wheat association mapping panel; IBWSN, international bread wheat screening nursery; IT, infection type; KALRO, Kenya Agricultural and Livestock Research Organization; LD, linkage disequilibrium; LR, leaf rust; LS, least squares; *Pt*, *Puccinia triticina; Pgt, Puccinia graminis; Pst, Puccinia striiformis*; QTL, quantitative trait loci; RKHS-M, reproducing kernel Hilbert spaces markers; RKHS-MP, reproducing kernel Hilbert spaces markers and pedigree; RKHS-P, reproducing kernel Hilbert spaces pedigree; RR-BLUP, ridge regression-best linear unbiased prediction; SR, stem rust; YR, stripe rust.

INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the major food crops in the world that is constantly threatened by several biotic stresses. Among the most significant fungal biotic stresses are the rusts that include leaf or brown rust (LR), stem or black rust (SR) and stripe or yellow rust (YR) caused by *Puccinia triticina* Eriks. (*Pt*), *Puccinia graminis* Pers. (*Pgt*) and *Puccinia striiformis* West. (*Pst*), respectively. Among these, LR is the most common rust that is globally distributed and can cause losses from 7 to 30% depending on the developmental stage (Roelfs et al., 1992; Marasas et al., 2004; Bolton et al., 2008; Huerta-Espino et al., 2011). Stem rust occurs mainly in warm weather

regions and can cause losses of up to 100% (Leonard and Szabo, 2005). Stripe rust occurs in cool, temperate regions and can cause yield losses ranging from 10 to 70% but up to 100% in highly susceptible cultivars (Chen, 2005). The most preferred management strategy for rusts is genetic resistance which is of two types, namely vertical and horizontal (Vanderplank, 1963). In a typical vertical resistance, the gene-for-gene interactions between the resistance genes of the host and the avirulence genes of the pathogen form the basis of resistance (Flor, 1956). Because of this incompatible interaction, hypersensitive cell death response is elicited. But, the major problem with this type of qualitative resistance is that it is ephemeral and can be easily overcome by the evolution of new virulent races of the pathogen. For example, the virulent stem rust race group Ug99 carries combined virulence to many genes deployed in the current wheat varieties and poses an enormous threat to global wheat production (Pretorius et al. 2000, Singh et al. 2015). Hence, many breeding efforts focus on horizontal, non-race specific, quantitative, slow rusting resistance which is the widely preferred mechanism to achieve durability, defined as the ability of a widely-deployed resistance gene to provide an economic level of protection over an extended period of time (Johnson, 1984). In this type of resistance, although the infection is not completely stopped, the spread of the disease is delayed and it is typically expressed in the adult plant stage (McIntosh et al., 1995). To date, about 76 LR resistance (*Lr*) genes, 59 SR resistance (*Sr*) genes, 76 YR resistance (*Yr*) genes and several quantitative trait loci (QTL) have been identified (McIntosh et al., 2016). Among these, the known race non-specific resistance genes are *Lr34/Yr18/Sr57*, *Lr46/Yr29/Sr58*, *Lr67/Yr46/Sr55*, *Lr68*, *Sr2/Lr27/Yr30* and *Yr36*.

Breeding for quantitative disease resistance is a challenge because of its complex inheritance and it is important to devise strategies for more effective evaluation and exploitation of this resistance. With this focus of accelerating breeding for quantitative resistance, one

promising approach that can potentially provide accurate predictions of the resistance phenotypes, enabling reduced time to parental selection and leading to increased genetic gain from selection is genomic selection (GS). Genomic selection uses dense genome wide markers to obtain the genomic estimated BVs of individuals (Meuwissen et al. 2001). It has been shown to be especially effective for improving quantitative traits, both in simulations (Bernardo and Yu 2007; Toosiet et al. 2010; Wong and Bernardo 2008) and in empirical studies (Crossa et al. 2010, 2014; Heslot et al. 2012; Lorenz et al. 2012; Ornella et al. 2012; Rutkoski et al. 2011, 2012, 2014). It uses a 'training population' comprising individuals that have been genotyped and phenotyped for traits of interest to generate BVs that can be used in selecting individuals for intermating in the next cycle of selection prior to phenotypic evaluation.

While some studies comparing prediction models have been reported (Lorenzana and Bernardo, 2009; Crossa et al., 2010; Heslot et al., 2012), our objective was to compare three genomic prediction models including genomic best linear unbiased prediction (GBLUP), GBLUP-A that was GBLUP with selected loci as fixed effects and reproducing kernel Hilbert spaces-markers (RKHS-M) with least squares (LS) approach, RKHS-pedigree (RKHS-P) and RKHS-markers and pedigree (RKHS-MP) to determine the BVs for seedling and/or adult plant resistance (APR) to LR, SR and YR. The GBLUP is a whole-genome regression approach that uses the genomic relationship matrix (G-matrix) calculated from markers instead of the pedigree relationship matrix. It has been successfully applied in the prediction of complex traits in humans, plants and animals (de Los Campos et al. 2013; Habier et al. 2013; VanRaden 2008; Yang et al. 2010). The RKHS semi-parametric approach for genomic prediction was proposed by Gianola (2006) and then by Gianola and van Kaam (2008) who argued that genomic interactions are much more complex than what could be handled by the standard parametric models. Several studies have

shown its effectiveness in genomic predictions (Crossa et al., 2010; De los Campos et al., 2010; Perez-Rodriguez et al., 2013). RKHS does not assume linearity and it is expected to capture some non-additive effects well. Since, the genetic architecture of seedling and APR to LR, SR and YR were different, we evaluated different models to determine which of them are appropriate for a given trait.

MATERIALS AND METHODS

Plant materials

For this study, we used the 45th and 46th international bread wheat screening nurseries (IBWSN) comprising 333 and 313 lines respectively. The IBWSNs are large screening nurseries that were initiated in 1967 and consist of 200 to 400 advanced lines from CIMMYT's (Centro Internacional de Mejoramiento de Maíz y Trigo) bread wheat breeding program (van Ginkel and Rajaram, 1993). These candidates were previously selected for biotic and abiotic stress resistance, grain yield and end-use industrial quality characteristics. They are evaluated in multiple trials in Mexico and cooperating locations globally. As such they are ideal for building prediction models as they are expected to have useful and novel genes for disease resistance with considerable variation in their BVs.

Disease evaluation and phenotypic data

Seedling evaluation for leaf rust and stripe rust

Seedling evaluations for LR (45th IBWSN – 2010 and 2012; 46th IBWSN - 2012) and YR (46th IBWSN - 2013) were conducted in CIMMYT's greenhouses at El Batan, Mexico. Rust inoculum was prepared by suspending freshly collected urediniospores (race MBJ/SP for *Pt* and

race Mex96.11 for *Pst*) in light mineral oil, Soltrol (Phillips 66 Co., Bartlesville, OK, USA). The plants were inoculated at the two-leaf stage, placed in a dew chamber overnight and then transferred to the greenhouse where the minimum, maximum, and average temperatures were 16.1°C, 30.0°C and 20.3°C. The LR seedling infection types (ITs) were recorded 10 days after inoculation using the 0 to 4 scale described in Roelfs et al. (1992). The responses were linearized to a 0-9 scale (; = 0, 0 = 0, 1- = 1, 1=2, 1+ = 3, 2- = 4, 2 = 5, 2+ = 6, 3- = 7, 3 = 8, 3+ = 9 and 4 = 9). For YR, the seedlings were incubated in a dew chamber at 7°C in the dark for 48 hours and then transferred to the greenhouse. The minimum, maximum, and average greenhouse temperatures were 6.3°C, 30.9°C and 17.3°C, respectively. YR infection types were recorded 14 days post-inoculation using a 0 to 9 scale as described by McNeal et al. (1971).

Adult plant response evaluation for leaf rust, stem rust and stripe rust

The 45th IBWSN entries were evaluated for APR to: LR at CIMMYT's headquarters, El Batan, Mexico during the 2010, 2012 and 2013 crop seasons; SR at Kenya Agricultural and Livestock Research Organization (KALRO), Njoro, Kenya during the 2010 and 2011 main seasons and YR at CIMMYT's research station, Toluca, Mexico during the 2011 and 2013 crop seasons. Similarly, the 46th IBWSN entries were evaluated for APR to LR at El Batan during the 2011 and 2013 crop seasons; SR at KALRO, Njoro during the 2011 main and off seasons and YR at Toluca during the 2011 and 2013 crop seasons, Quito, Ecuador during the 2012 crop season and KALRO, Njoro during the 2011 main season. The modified Cobb Scale (Peterson et al. 1948) was used to score rust severity at the adult plant stage to determine the percentage of infected tissue (0-100%). Evaluations were conducted at three time points between early and late dough stages. The first evaluation was done when the severity of susceptible check (Avocet) reached 80% followed by

two more evaluations at weekly intervals. For all the rust evaluations, the lines were sown in 0.7-m long paired rows on top of 30-cm-wide raised beds. For LR, a mixture of the susceptible genotypes 'Avocet+*Yr24*' and 'Avocet+*Yr26'* was planted as spreader rows around the experimental field. The spreader rows and hills were artificially inoculated with urediniospores of the two prevalent Mexican *Pt* races, MBJ/SP and MCJ/SP suspended in Soltrol oil to initiate an epidemic. These two races differ by their virulence to the *Lr26* gene (MBJ/SP has partial virulence for *Lr26*, while MCJ/SP has complete virulence). The inoculations were carried out twice when the plants were at the 6-leaf stage. For SR evaluation, a border row of spreaders was planted surrounding the field and sprayed twice with fresh urediniospores of *Pgt* race TTKST suspended in Soltrol to create an artificial rust epidemic. The plants within the border rows were inoculated by injecting a suspension of freshly collected urediniospores in water using a hypodermic syringe, twice prior to booting (growth stage Z35-Z37) (Zadoks et al. 1974). For YR evaluation, spreaders consisted of a mixture of six susceptible wheat lines derived from an Avocet/Attila cross. The 4-week old spreaders and hills were inoculated three times, at three to four day intervals with mixed *Pst* isolates, Mex96.11 and Mex08.13. While Mex96.11 is virulent to *Yr27* and avirulent to *Yr31*, it is the reverse for Mex08.13.

The phenotypic distributions for all the diseases were transformed using the boxcox (Box and Cox, 1964) function in the 'R' statistical program.

Genotyping

The nurseries were genotyped using the genotyping-by-sequencing (GBS) method to obtain dense genome-wide coverage (Elshire et al. 2011). GBS markers were obtained using the method described by Poland et al. (2012). After filtering for markers with missing data greater

than 50%, minor allele frequency less than 10% and pairwise marker correlation (r^2) greater than 0.95 (for redundancy), 5102 markers for the 45th IBWSN and 8066 markers for the 46th IBWSN were obtained. Missing data was imputed using the expectation-maximization algorithm implemented in the rrBLUP software package (Endelman, 2011). The lines were also filtered for missing data greater than 50% which resulted in 267 and 305 lines in the 45th and 46th IBWSN, respectively.

Relationship matrix and heritability estimation

The G-matrix was calculated according to VanRaden (2008) and implemented in the 'R' package rrBLUP (Endelman, 2011). The relationship matrix was centered and standardized for all the analyses. Heritability was calculated on a line mean basis and estimates of the genetic and residual variances were obtained using the average information-restricted maximum likelihood algorithm (Gilmour et al. 1995) implemented in the 'heritability' package in R (Kruijer et al., 2015).

Prediction models

Least squares (LS)

A step-wise least squares (LS) approach was used which involves an initial marker ranking and selection step. First, genome wide association analysis was conducted in the training set to calculate marker p-values. Then the markers were ranked according to their p-values for variable selection. For each iteration i through j, a marker was added to the model, starting from the marker with the lowest p-value,

$$\boldsymbol{y} = 1_n\boldsymbol{\mu} + \boldsymbol{X}_i\beta_i \ldots\ldots \boldsymbol{X}_j\beta_j + \boldsymbol{\varepsilon}$$

where **y** is the phenotype, μ is the mean, β_i denotes the effect of the i^{th} marker and $\mathbf{X}_i$ denotes the i^{th} marker's genotype matrix. The 5-fold cross validation accuracy was calculated within the training set after each iteration and the model with j-1 markers was selected when the $Accuracy_{j-1}$>$Accuracy_j$. The second step involved marker effects estimation from the selected model that was then used to predict the BVs of the individuals. To obtain the chromosomal locations of the significant markers, the basic local alignment search tool (BLAST) (https://triticeaetoolbox.org/wheat/viroblast/viroblast.php) in the Triticeae toolbox website was used. A nucleotide BLAST (BLAST-n) was performed against the wheat markers in T3 database (updated on April 2015); wheat contigs (1A to 7D) from the wheat CSS genome reference v2, September 2014 and wheat chromosomes (1A to 7D) and unsorted scaffolds from the IWGSC1.0+popseq (November 2014) (Chapman et al. 2015). This approach would help to identify markers that are similar in other populations genotyped by GBS and also enable us to compare across studies using marker synonyms.

Genomic best linear unbiased prediction (GBLUP) and GBLUP with selected loci as fixed effects (GBLUP A)

For GBLUP, the BVs of individuals was predicted using the mixed model,

$$\boldsymbol{y}=\boldsymbol{1}_n\boldsymbol{\mu}+\boldsymbol{Zu}+\boldsymbol{\varepsilon}$$

where **y** is the vector of the response phenotypic trait, **μ** is the overall mean vector, **u** is the vector of genotype effects that are assumed to be multivariate normal random effects ($\mathbf{u} \sim N(\mathbf{0}, \mathbf{G}\sigma^2_u)$), **Z** is the design matrix for the random effects and **ε** is the vector of independent residuals assumed to have a multivariate normal distribution ($\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^2_e)$). The 'R' package, rrBLUP (Endelman,

2011) was used to implement GBLUP. We also evaluated a GBLUP A model with selected loci modeled as fixed effects,

$$\boldsymbol{y}=\boldsymbol{1}_n\boldsymbol{\mu}+\boldsymbol{X}_i\beta_i\ldots\ldots\boldsymbol{X}_j\beta_j+\boldsymbol{Zu}+\boldsymbol{\varepsilon}$$

Reproducing kernel Hilbert spaces (RKHS)

The RKHS model using a Gaussian kernel is of the form:

$$\boldsymbol{y}_i=\boldsymbol{w}_i'\boldsymbol{\beta}+\boldsymbol{z}_i'\boldsymbol{u}+\sum_{j=1}^{n}exp\left[-\frac{(x_i-x_j)'(x_i-x_j)}{h}\right]\alpha_j+\boldsymbol{\varepsilon_i}$$

where x_i and x_j are the observed marker genotypes of individuals, $\mathbf{w_i}$ and $\mathbf{z_i}$ are the incidence vectors, β is the vector of location effects, **u** is the vector of additive genetic effects, α_j is the regression coefficient and $\boldsymbol{\varepsilon_i}$ is the error term [$\varepsilon_i \sim N(0, \mathbf{I}\sigma^2_e)$] (Gianola et al., 2006). The additive genetic effects $\mathbf{u} \sim N(0, \mathbf{K}\sigma^2_g)$, where K is the reproducing Gaussian kernel, $K(x_i,x_j) = exp\left[-\frac{(x_i-x_j)'(x_i-x_j)}{h}\right]$ and 'h' is the bandwidth parameter. We implemented three RKHS models in the BGLR package (Pérez and de Los Campos, 2014) namely (i) RKHS markers (RKHS-M) using the G-matrix calculated from markers (ii) RKHS pedigree (RKHS-P) using the pedigree relationship matrix which was obtained from the pedigree and was twice the co-efficient of ancestry (iii) RKHS markers and pedigree (RKHS-MP) where part of the additive effect was captured by regression on the markers and also with the (co)variance relationship derived from the pedigree. We fitted these models with three arbitrarily chosen bandwidth parameters and then averaged the three accuracies.

Prediction accuracies

The predictive ability of the models was assessed using the Pearson's correlation between the observed and the cross-validated estimated BVs which is the prediction accuracy. We used the 10-fold cross-validation where the whole dataset was divided into 10 folds and nine of them (240 lines and 275 lines in the 45th and 46th IBWSN, respectively) were used as a training set to estimate the marker effects which were then used to predict the BVs in the 10th fold, referred to as the validation set (27 lines and 30 lines in the 45th and 46th IBWSN, respectively).

RESULTS

Phenotypic data analysis

The phenotypic distributions of the rusts in the 45th and 46th IBWSN are shown in Figure 1.1. In both trials, the average correlation between LR seedling resistance and APR was very low (0.1 and 0.3 for the 45th and 46th IBWSN, respectively) indicating that the genetic bases of seedling resistance and APR were different.

Relationship and heritability analysis

Heatmap of the genomic and the pedigree based relationship matrices for the 45th and 46th IBWSN (Figure 1.2) indicated that the lines in the 46th IBWSN had a slightly higher relationship among them than those in the 45th IBWSN. The 267 lines in the 45th IBWSN comprised one family with eight full-sibs, one with six full-sibs, one with five full-sibs, seven with four full-sibs, 15 with three full-sibs, 37 with two full-sibs and 101 crosses represented by one individual per cross. The 305 lines in the 46th IBWSN comprised one family with seven full-sibs, two with six full-sibs, seven with four full-sibs, 12 with three full-sibs, 34 with two full-sibs and 154 with one individual

per cross. We also observed that the pedigree relationship matrices for both nurseries indicated a higher relationship among the lines than the marker-based matrices because it does not account for Mendelian sampling.

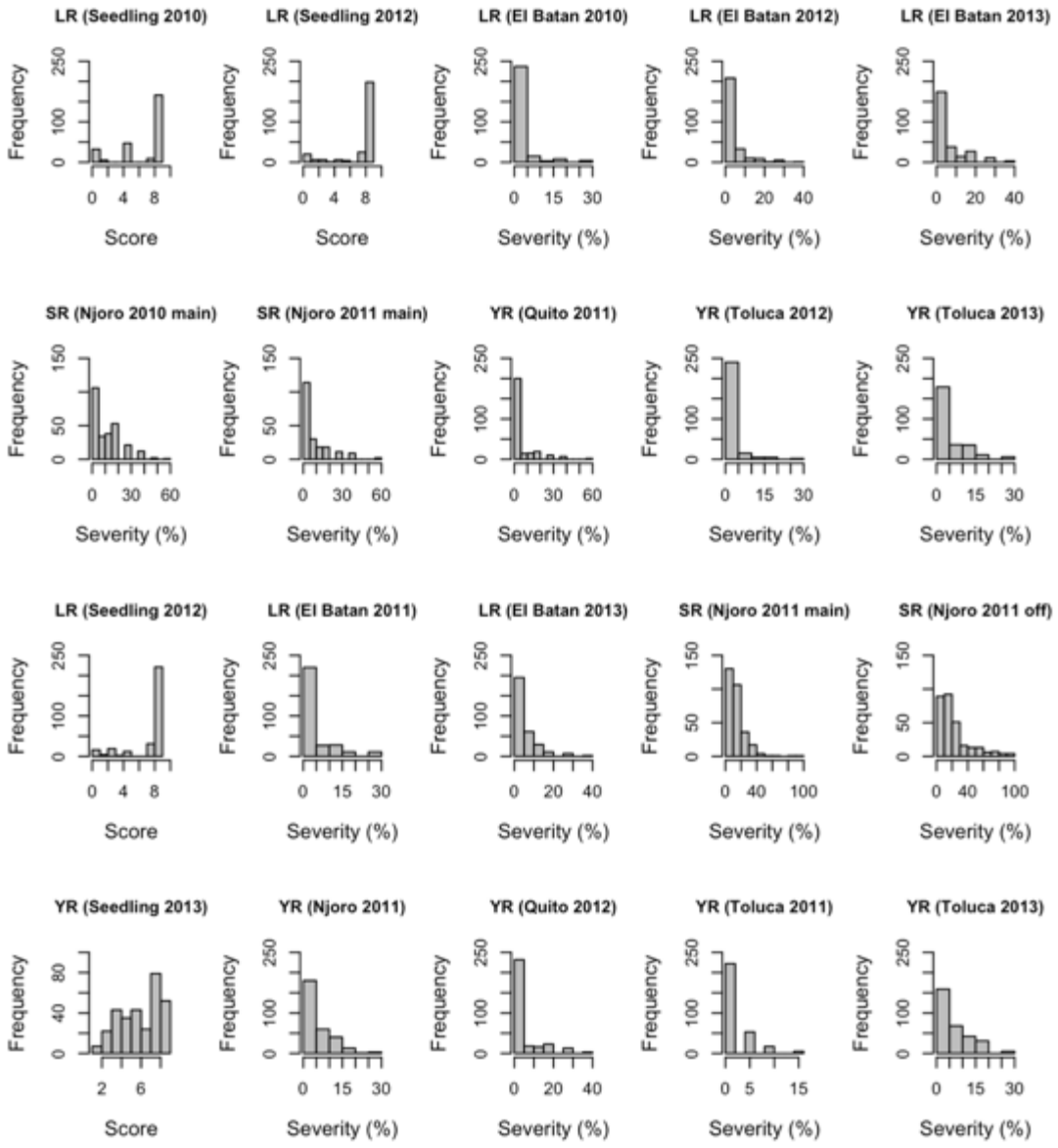

Figure 1.1: Phenotypic distributions for leaf rust (LR), stem rust (SR) and stripe rust (YR) in the 45th (top two panels) and 46th (lower two panels) international bread wheat screening nurseries (IBWSN)

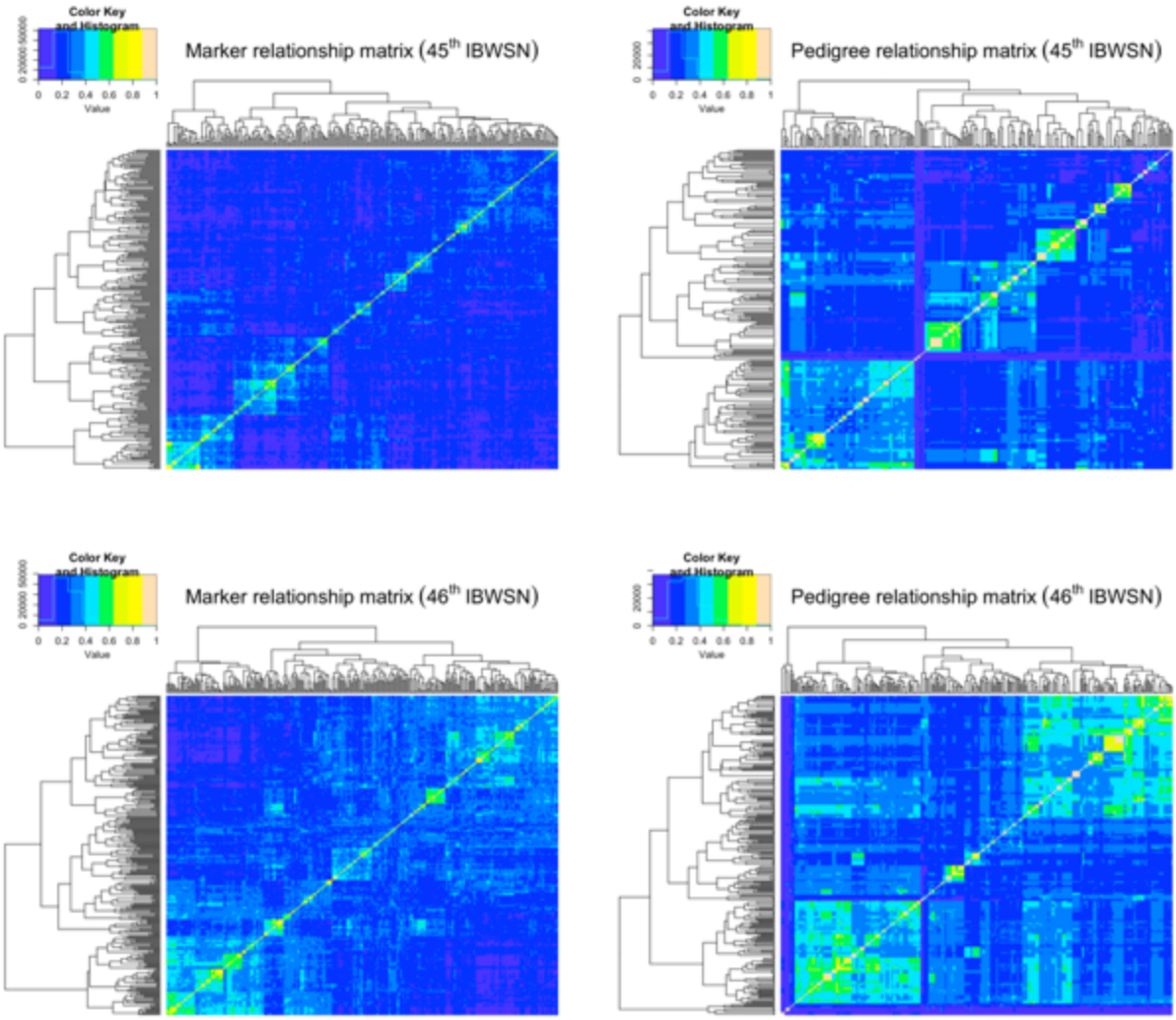

Figure 1.2: Heat map of the marker and pedigree based relationship matrices for the 45th and 46th international bread wheat screening nurseries (IBWSN) illustrating the familial relatedness (kinship) between the individuals

In the 45th IBWSN, the broad-sense line-mean heritability was the highest for LR seedling (0.72) followed by SR APR (0.59), LR APR (0.58) and YR APR (0.26). In the 46th IBWSN, the highest heritability was obtained for LR APR (0.6), followed by SR APR (0.5) and YR APR (0.48). The broad sense heritability was very high for LR seedling (0.87) and YR seedling (0.86).

Markers significantly associated with leaf, stem and stripe rust resistance

The markers that were significantly associated with LR, SR and YR resistance in the 45th and 46th IBWSN and used as fixed effects in the LS model are shown in Tables 1.1 and 1.2, respectively. Only the markers that perfectly matched with a marker in the T3 database and were significant in at least five folds are reported. The BLAST results for all the markers and other synonyms are reported in Supplementary Table 1. For LR seedling resistance in the 45th IBWSN, marker GBS_24751 (0cM) on chromosome 2BS explained the highest variation (18%) in the 2010 dataset. Marker GBS_37247 on chromosome 1DS was significant in both the replications and explained 15 and 24% of the average variation. In the 46th IBWSN, the marker GBS_19971 on chromosome 1DS was significant in all the folds and explained an average of 33% of the variation for LR seedling resistance. In addition, a marker on chromosome 3B and another marker on chromosome 1DS were also significant. For LR APR in the 45th IBWSN, the marker GBS_30281 on chromosome 4AL was significant in all the three datasets and explained 8 to 12% of the average variation. The only other significant marker with known position was GBS_8842 on chromosome 3AS in the 2010 dataset. In the 46th IBWSN, GBS_40747 on chromosome 2D was significantly associated with LR APR in the El Batan 2011 dataset and explained 10% of the average variation. Also, markers GBS_18425 and GBS_2400 both on chromosome 3AS and GBS_1491 on chromosome 3AL were significant in the El Batan 2013 dataset.

For SR APR, in the 45th IBWSN, the marker GBS_22856 on chromosome 3B was significantly associated and explained an average variation of 16 and 18% in the Njoro 2010 and 2011 datasets, respectively. In addition, markers GBS_36529 on chromosome 3B and GBS_2454 on chromosome 5B were significant in the Njoro 2010 dataset and marker GBS_13047 on chromosome 3B was significant in the Njoro 2011 dataset. In the 46th IBWSN, GBS_23856 on

chromosome 1AL and GBS_1505 on chromosome 3B were significant in the Njoro 2011 main season. Markers, GBS_28025 on chromosome 6BS, GBS_1505 on chromosome 3B and GBS_20060 on chromosome 6DS were significant in the 2011 off season. For YR seedling (2013) in the 46th IBWSN, the marker GBS_702 on chromosome 2AS was significant in all the folds and explained an average 58.5% of the variation. For YR APR, the marker GBS_6432 on chromosome 2AS was significant in all the datasets and explained an average of 16 to 32% of the variation. GBS_702 on chromosome 2AS was significant in all the folds in the Toluca 2013 dataset and explained 29% of the average variation. In the 46th IBWSN, the marker, GBS_702 on chromosome 2AS, was significant in all the folds in all the YR APR datasets and explained an average variation of 26 to 41%.

Table 1.1: Quantitative trait loci (QTL) linked markers used as fixed effects in the least-squares (LS) model for the 45th international bread wheat screening nursery (IBWSN)

	Dataset	Marker	Marker synonym[a]	Chromosome	Genetic position (Popseq map)	Physical position (Popseq map)	Expected genes	Average p-value	Average R^2	Frequency[b]
Leaf rust	Seedling 2010	GBS_24751	gbsHWWAMP38350	2BS	0	3863480	*Lr16*	2.15E-10	0.18	0.8
		GBS_37247	WCSS1_contig470290_1DS-434	1DS	2.7	1241625	*Lr42*	2.57E-09	0.15	0.5
	Seedling 2012	GBS_37247	WCSS1_contig470290_1DS-434	1DS	2.7	1241625	*Lr42*	4.33E-14	0.24	1
		GBS_38357	-	-	-	-	-	2.04E-10	0.17	0.8
	El Batan 2010	GBS_8842	WCSS1_contig3334901_3AS-4592	3AS	9.4	3823808	-	5.24E-07	0.12	1
		GBS_30281	WCSS1_contig7120458_4AL-2167	4AL	121.9	210542445	-	2.13E-07	0.12	0.9
	El Batan 2012	GBS_12317	-	-	-		-	3.42E-06	0.09	0.5
		GBS_30281	WCSS1_contig7120458_4AL-2167	4AL	121.9	210542445	-	6.67E-06	0.08	0.5
	El Batan 2013	GBS_1926	-	-	-	-	-	5.60E-08	0.12	0.9
		GBS_30281	WCSS1_contig7120458_4AL-2167	4AL	121.9	210542445	-	6.21E-08	0.12	0.9
		GBS_5135	-	-	-	-	-	2.69E-07	0.11	0.5
Stem rust	Njoro 2010	GBS_22856	WCSS1_contig10511286_3B-5360	3B	-	90941978	*Sr2*	3.89E-10	0.16	1
		GBS_36529	WCSS1_contig10759567_3B-1965	3B	76.4	370963380	*Sr12*	9.17E-09	0.14	0.5
		GBS_2454	WCSS1_contig2284473_5BS-12686	5BS	4.2	4471954	-	3.19E-09	0.14	0.5
	Njoro 2011	GBS_22856	WCSS1_contig10511286_3B-5360	3B	-	90941978	*Sr2*	1.04E-08	0.18	1
		GBS_13047	gbsHWWAMP18106	3B	0	312463	*Sr2*	8.10E-08	0.17	0.9
		GBS_23598	-	-	-	-		1.20E-07	0.15	0.7
Stripe rust	Quito 2011	GBS_6432	WCSS1_contig5219749_2AS-4945	2AS	8.8	6792755	*Yr17*	2.91E-10	0.16	1
	Toluca 2012							2.41E-13	0.21	1
	Toluca 2013	GBS_702	WCSS1_contig5304580_2AS-10182	2AS	0	944474		0.0E+00	0.29	1
		GBS_6432	WCSS1_contig5219749_2AS-4945	2AS	8.8	6792755		0.0E+00	0.32	0.9

[a] Markers prefixed by gbsHWWAMP are from the hard winter wheat association mapping panel available in T3 and markers prefixed by WCSS1_contig are from the CSS GBS 2014 physical map, where 'WCSS1' stands for wheat chromosome survey sequence.

[b] the frequency of the marker in the ten cross-validation folds.

Table 1.2: Quantitative trait loci (QTL) linked markers used as fixed effects in the least-squares (LS) model for the 46th international bread wheat screening nursery (IBWSN)

Dataset		Marker	Marker synonym[a]	Chromosome	Genetic position (Popseq map)	Physical position (Popseq map)	Expected genes	Average p-value	Average R^2	Frequency[b]
Leaf rust	Seedling 2012	GBS_19971	WCSS1_contig1905752_1DS-4628	1DS	5.4	2073708	*Lr42*	<E-16	0.33	1
		GBS_28186	gbsHWWAMP44231	3B	25.3	15366258	-	<E-16	0.26	1
		GBS_28376	WCSS1_contig1898017_1DS-2235	1DS	11.0	3306922	*Lr42*	<E-16	0.39	0.9
	El Batan 2011	GBS_40747	gbsHWWAMP55344	2D	17.3	8198944	-	3.08E-07	0.10	0.8
		GBS_38496	-	-	-	-	-	2.27E-07	0.10	0.7
	El Batan 2013	GBS_18425	WCSS1_contig3419689_3AS-1090	3AS	60.6	61387121	-	5.08E-10	0.14	1
		GBS_2400	WCSS1_contig3361063_3AS-2705	3AS	53.4	14901099	-	2.83E-10	0.13	0.6
		GBS_1491	gbsHWWAMP1393	3AL	63.1	133091112	-	3.70E-10	0.14	0.5
Stem rust	Njoro 2011 main	GBS_23856	gbsHWWAMP37196	1AL	86.5	220028370	-	3.44E-08	0.12	0.9
		GBS_1505	-	3B	-	91939	*Sr2*	6.79E-08	0.11	0.5
	Njoro 2011 off	GBS_28025	WCSS1_contig3042477_6BS-5453	6BS	65.1	70672093	-	1.18E-08	0.13	1
		GBS_1505	-	3B	-	91939	*Sr2*	1.03E-07	0.11	0.6
		GBS_20060	WCSS1_contig2078323_6DS-22086	6DS	2.5	2067639	*SrTmp*	1.33E-08	0.12	0.5
Stripe rust	Seedling 2013	GBS_702	WCSS1_contig5304580_2AS-10182	2AS	0	944474	*Yr17*	<E-16	0.58	1
	Quito 2012							<E-16	0.26	1
	Njoro 2011							<E-16	0.27	1
	Toluca 2011							<E-16	0.41	1
	Toluca 2013							<E-16	0.41	1

[a] Markers prefixed by gbsHWWAMP are from the hard winter wheat association mapping panel available in T3 and markers prefixed by WCSS1_contig are from the CSS GBS 2014 physical map, where 'WCSS1' stands for wheat chromosome survey sequence.

[b] the frequency of the marker in the ten cross-validation folds.

Prediction accuracies

Prediction accuracies for LR, SR and YR resistance in the 45th and 46th IBWSNs are shown in Table 1.3.

Prediction accuracies for leaf rust seedling and adult plant resistance

For LR seedling in the 45th IBWSN, the highest prediction accuracy was obtained using the RKHS-MP and RKHS-P, respectively in the 2010 and 2012 datasets. The lowest accuracy was obtained using the LS approach and GBLUP resulted in 125.8% and 38.1% increase in accuracy over LS in the two datasets. While RKHS-P model performed similar to the other genome-wide models in the 2010 dataset, it gave a 23.7% increase in accuracy over the RKHS-M in the 2012 dataset. There were no significant differences in the accuracies obtained from GBLUP, GBLUP A and RKHS-M. In the 46th IBWSN, the highest accuracy for LR seedling resistance was obtained using the GBLUP A followed by the RKHS-MP, LS, RKHS-M and GBLUP which gave similar accuracies. The RKHS-P yielded the lowest prediction accuracy, but it was only 6.55% lower than RKHS-M. For LR APR, it was observed that RKHS-MP gave the highest accuracies and LS, the lowest in all five datasets. The increase in accuracy obtained from using GBLUP over LS varied across the different datasets and ranged from 26.5 to 241.7%. GBLUP A performed similar to the GBLUP in all the datasets except in the El Batan 2012 dataset (45th IBWSN), where the fixed effect markers explained very little variation. The accuracies obtained using pedigree and genome-wide marker based models were not significantly different in all the datasets but there was a slight increase in accuracy using genome-wide markers in the El Batan 2012 (45th IBWSN) and El Batan 2013 (46th IBWSN) datasets (20.6% and 10.4% respectively). GBLUP and RKHS-M gave similar accuracies in all the datasets.

Prediction accuracies for stem rust adult plant resistance

For SR APR, the lowest prediction accuracy in all four datasets was obtained using LS and GBLUP resulted in 43.9% to 74.2% increase in accuracy over LS. The highest accuracy was obtained with RKHS-MP in two datasets, GBLUP A in one dataset and with both GBLUP and RKHS-M in the other dataset. The RKHS-P performed similar to the GBLUP in one dataset and slightly better than GBLUP (6.8% increase in accuracy) in another dataset. But we observed a decrease in accuracy of 10.2% and 27.7% using the RKHS-P vs RKHS-M in two datasets. As observed for LR, GBLUP and RKHS-M gave similar accuracies in all the datasets.

Prediction accuracies for stripe rust seedling and adult plant resistance

For YR seedling resistance in the 46th IBWSN, the highest accuracies were obtained using GBLUP A followed by LS, RKHS-MP, GBLUP, RKHS-M and RKHS-P models. Although, RKHS-P gave the lowest accuracy, the increase in accuracy using RKHS-M over the pedigree was only 4.3%. Least squares performed slightly better than the GBLUP and resulted in 5.5% increase in accuracy. For YR APR, in the 45th IBWSN, the highest accuracy was obtained with GBLUP A in the Quito 2011 dataset, with LS and GBLUP A in the Toluca 2012 dataset and with GBLUP A and RKHS-MP in the Toluca 2013 dataset. Least squares performed similar to the GBLUP in the Quito 2011 dataset, slightly better than the GBLUP in the Toluca 2012 dataset (15.4% increase in accuracy) and poorer than the GBLUP (20.3% decrease in accuracy) in the Toluca 2013 dataset. Although, the RKHS-P model gave the lowest accuracies in two datasets, the increase in accuracy using markers was not significant (ranged from 3% to 11.8%). The GBLUP, RKHS-M and RKHS-MP models gave similar accuracies in all the datasets. In the 46th IBWSN, RKHS-MP gave the

highest accuracy in the Quito 2012 dataset; RKHS-MP and GBLUP A in the Njoro 2011 dataset and GBLUP A in the Toluca 2011 and 2013 datasets. The RKHS-P model performed similar to RKHS-M in all the datasets, except the Toluca 2013 dataset where RKHS-M resulted in 23.6% increase in accuracy over the RKHS-P. We also observed that LS performed similar to GBLUP in the Njoro 2011 dataset, slightly better than GBLUP in the Toluca 2011 dataset (5% increase in accuracy) and slightly poorer than GBLUP in the Quito 2012 and Toluca 2013 datasets (7.3% and 7.4% decrease in accuracy). GBLUP and RKHS-M models yielded similar accuracies in all the datasets.

Table 1.3: Prediction accuracies for leaf rust (LR), stem rust (SR) and stripe rust (YR) resistance in the 45th and 46th international bread wheat screening nurseries (IBWSN)

Trait	Dataset	IBWSN	LS	GBLUP	GBLUP A	RKHS-M	RKHS-P	RKHS-MP
Leaf rust	Seedling 2010	45th	0.31 ± 0.09	0.7 ± 0.03	0.69 ± 0.03	0.71 ± 0.03	0.7 ± 0.03	**0.74 ± 0.03**
	Seedling 2012	45th	0.42 ± 0.1	0.58 ± 0.05	0.6 ± 0.05	0.59 ± 0.05	**0.73 ± 0.05**	0.72 ± 0.05
	Seedling 2012	46th	0.66 ± 0.04	0.64 ± 0.05	**0.7 ± 0.03**	0.65 ± 0.05	0.61 ± 0.07	0.67 ± 0.05
	El Batan 2010	45th	0.34 ± 0.05	0.43 ± 0.05	0.43 ± 0.04	0.43 ± 0.05	0.42 ± 0.07	**0.46 ± 0.05**
	El Batan 2012	45th	0.12 ± 0.07	0.41 ± 0.05	0.26 ± 0.07	0.41 ± 0.05	0.34 ± 0.06	**0.41 ± 0.05**
	El Batan 2013	45th	0.29 ± 0.06	0.47 ± 0.06	0.44 ± 0.06	0.48 ± 0.06	0.5 ± 0.06	**0.52 ± 0.06**
	El Batan 2011	46th	0.28 ± 0.05	0.51 ± 0.04	0.49 ± 0.05	0.51 ± 0.04	0.5 ± 0.03	**0.53 ± 0.04**
	El Batan 2013	46th	0.38 ± 0.03	0.52 ± 0.04	0.51 ± 0.03	0.53 ± 0.03	0.48 ± 0.03	**0.56 ± 0.03**
Stem rust	Njoro 2010 main	45th	0.41 ± 0.05	0.59 ± 0.04	0.64 ± 0.03	0.59 ± 0.04	0.63 ± 0.03	**0.65 ± 0.03**
	Njoro 2011 main	45th	0.41 ± 0.08	0.59 ± 0.05	**0.62 ± 0.04**	0.59 ± 0.05	0.53 ± 0.07	0.58 ± 0.06
	Njoro 2011 main	46th	0.31 ± 0.04	0.54 ± 0.05	0.54 ± 0.05	0.54 ± 0.05	0.55 ± 0.06	**0.62 ± 0.05**
	Njoro 2011 off	46th	0.31 ± 0.03	**0.47 ± 0.06**	0.43 ± 0.05	**0.47 ± 0.06**	0.34 ± 0.04	0.45 ± 0.06
Stripe rust	Seedling 2013	46th	0.77 ± 0.03	0.73 ± 0.03	**0.78 ± 0.02**	0.73 ± 0.03	0.7 ± 0.03	0.74 ± 0.03
	Quito 2011	45th	0.37 ± 0.05	0.39 ± 0.06	**0.41 ± 0.06**	0.38 ± 0.07	0.34 ± 0.08	0.39 ± 0.07
	Toluca 2012	45th	**0.45 ± 0.04**	0.39 ± 0.04	**0.45 ± 0.04**	0.39 ± 0.05	0.37 ± 0.04	0.39 ± 0.04
	Toluca 2013	45th	0.55 ± 0.03	0.69 ± 0.02	**0.7 ± 0.02**	0.68 ± 0.02	0.66 ± 0.03	**0.7 ± 0.03**
	Quito 2012	46th	0.51 ± 0.03	0.55 ± 0.03	0.6 ± 0.03	0.54 ± 0.03	0.58 ± 0.03	**0.61 ± 0.03**
	Njoro 2011	46th	0.51 ± 0.03	0.52 ± 0.03	**0.56 ± 0.03**	0.52 ± 0.04	0.55 ± 0.04	**0.56 ± 0.04**
	Toluca 2011	46th	0.63 ± 0.03	0.6 ± 0.03	**0.65 ± 0.03**	0.59 ± 0.02	0.64 ± 0.02	0.63 ± 0.02

	Toluca 2013	46th	0.63 ± 0.04	0.68 ± 0.03	**0.71 ± 0.04**	0.68 ± 0.03	0.55 ± 0.06	0.66 ± 0.04

IBWSN, International bread wheat screening nursery; LS, least squares; GBLUP, genomic best linear unbiased prediction; GBLUP A, genomic-BLUP with selected loci as fixed effects; RKHS-M, reproducing kernel Hilbert spaces markers; RKHS-P, reproducing kernel Hilbert spaces pedigree; RKHS-MP, reproducing kernel Hilbert spaces markers and pedigree.

DISCUSSION

Genomic prediction for LR seedling resistance resulted in an 82% average increase in accuracy over LS in the 45th IBWSN. But, LS performed similar to genome-wide marker models in the 46th IBWSN. This can be attributed to the two significant markers on chromosome 1DS (5.4cM and 11cM) that were used as fixed effects and explained a large amount of the variability in the folds (33% and 39%). In this case, genome-wide markers would not be required for high accuracy, suggesting that the genetic architecture of resistance in a given population is an important factor that determines the appropriate model. We believe these markers to be linked to the *Lr42* seedling resistance gene based on their distal location in the chromosome and also the presence of this gene in Quaiu (Basnet et al., 2013), which was used as a parent for several crosses. A marker at about the same position on chromosome 1DS (2.7cM) was also significant in both the datasets in the 45th IBWSN. While it could also be linked to the *Lr42* gene, it explained only 15 to 24% of the variability in this nursery which resulted in lower accuracies using the LS. There was also another marker at the distal end of chromosome 2BS (0 cM) used as a fixed effect in the 2010 dataset, which is likely to be linked to *Lr16,* a seedling effective race-specific resistance gene. *Lr16* is present at a high frequency in CIMMYT germplasm, especially in lines derived from Waxwing and Francolin parentage (Lan et al., 2014), which were used as parents in several crosses.

Genomic prediction for APR to LR and SR, yielded an average increase in accuracy of 89.8% and 53.4% respectively, over LS in both the nurseries. Because LR and SR APR had moderate heritabilities and are quantitative traits conditioned by many genes with small effects, the poor performance of the LS was expected. For LR APR in the 45th IBWSN, the marker on chromosome 4AL that was significant in all the datasets did not coincide with any of the known genes which are effective to this *Pt* race and may be identifying a novel QTL. A marker on

chromosome 3AS (9.4cM) was significant in all the folds in the 2010 dataset. Although, *Lr63* is the only known gene mapped to the distal end of chromosome 3AS (Kolmer et al., 2010), it is unlikely that it is present in these lines considering its origin. In the 46th IBWSN, a marker on chromosome 2D (17.3cM) was significant in the 2011 dataset and three markers on chromosome 3A (53.4cM to 63.1cM) were significant in the 2013 dataset. Since, their positions could not be compared to any of the known genes in these chromosomes and the catalogued genes are not effective to this *Pt* race, they might be identifying novel QTL. Stem rust APR in the 45th IBWSN was associated with markers at two locations on chromosome 3B (0cM and 76.4cM). The marker at the distal end of chromosome 3B might be linked to the durable stem rust resistance gene, *Sr2* which is present in a high frequency in CIMMYT lines. The other marker on chromosome 3B might be linked to the *Sr12* resistance gene, which despite being ineffective against Ug99 alone, was suggested to confer APR in combination with other resistance loci by complementary epistasis (Rouse et al., 2014). XwPt6047, the marker closely linked to the *Sr12* gene (Rouse et al., 2014) is located at 52.7cM in the CIMMYT integrated DArT map (Crossa et al., 2007), but it was not possible to obtain its relative position in the popseq map. In addition to the markers on chromosome 3B, a marker at the distal end of chromosome 5BS (4.2cM) was also significant. While it is was not possible to determine what gene it was linked to, a minor QTL for Ug99 resistance has been reported on the distal end of chromosome 5BS by Yu et al. (2011). In the 46th IBWSN, SR APR was associated with a marker on an unknown location on chromosome 3B in both seasons and one marker each on chromosome 1AL (86.5cM), chromosome 6BS (65.1cM) and chromosome 6DS (2.5cM). The position of the markers on chromosomes 1AL and 6BS could not be compared to previously reported Ug99 resistance QTL as relative markers were not

available. We believe that the marker on chromosome 6DS is linked to the *SrTmp* gene, but it is no longer effective against Ug99 (Newcomb et al., 2016).

For seedling resistance to YR, we observed that the GBLUP A and LS performed slightly better than the GBLUP. This can be attributed to the very high heritability of the trait and the marker, GBS_702 on chromosome 2AS that explained a large variation in the folds. This is another case where genomic prediction is not necessary for high accuracy. For YR APR in the 45th IBWSN, GBLUP A performed the best in all the datasets and LS also performed well except in one dataset. This is due to markers GBS_6432 and GBS_702 on chromosome 2AS that explained a large variation. Similarly, in the 46th IBWSN, the GBLUP A model had the highest accuracy in most datasets and the high accuracies obtained from both LS and GBLUP A were due to the marker, GBS_702 on chromosome 2AS that had a large effect. Unlike LR and SR, APR to YR in these nurseries behaved as a simple trait and could be predicted well using LS. The significant association of the same marker to both seedling resistance and APR indicates that it is an all-stage resistance gene that we believe to be *Yr17* or a closely linked gene. The *Yr17* gene is located at the distal end of chromosome 2AS which is also the location of GBS_702 (0 cM) and GBS_6432 (8.8cM). Although, *Yr17* is closely linked to *Lr37* and *Sr38,* it is to be noted that races MBJ/SP and MCJ/SP are virulent to *Lr37* and the Ug99 group of races in Kenya are virulent to *Sr38*.

Overall, our prediction results indicate that genome-wide marker based prediction models were more accurate than LS in most datasets, which is consistent with several previous studies (Meuwissen et al., 2001; Bernardo and Yu, 2007; Habier et al., 2007; Muir, 2007; Piyasatian et al., 2007; Lorenzana and Bernardo, 2009; Moser et al., 2009; Heffner et al., 2011a; b, Rutkoski et al., 2014, 2012). We obtained an average of 42% increase in accuracy using the GBLUP compared to LS. This is comparable to previous reports: Meuwissen et al. (2001) obtained a 41% greater

accuracy using RR-BLUP than stepwise regression in simulations; Bernardo and Yu (2007) obtained an 18 and 43% improvement in the responses using GS compared to marker assisted recurrent selection in their simulation study for a trait that has high and low heritability; Piyasatian et al. (2007) obtained a 32% increase in accuracy using RR over stepwise regression in earlier generations; Heffner et al. (2011a) reported 28% higher average accuracies using GS than marker-assisted selection in a population of advanced cycle winter wheat breeding lines. The poor predictive ability of LS for some traits results from the fact that complex traits are controlled by many QTL, thereby supporting the infinitesimal model of Fisher (1918) and the use of single-QTL models is naïve (Dekkers and Hospital 2002; Gianola 2006; Meuwissen et al. 2001). We also observed that when the trait was controlled by large effect loci, the benefits of GS over genomic prediction models was low. This was the case for seedling resistance to LR and YR in the 46th IBWSN and also APR to YR in several datasets in both nurseries. There were also some datasets in our study where the LS performed slightly better than the GBLUP. This can be attributed to the fact that LS may better capture large effect QTL and eliminate the noise due to the markers with near zero effect that are included in the GBLUP. Hence, we would recommend using the LS for oligogenic resistance and GBLUP for quantitative resistance. The GBLUP A performed well for traits where the fixed effect markers explained a large amount of the variation. A previous study by Rutkoski et al. (2014) for quantitative APR to SR in wheat reported that the GBLUP A had higher accuracy than GBLUP alone. Although, the average increase in accuracy using GBLUP A over GBLUP was only 1.3% in our study, it ranged between 15.4 to -36.6%.

Our results also indicate that the RKHS-M model performed similar to the GBLUP, although several studies have reported that non-parametric models performed better than the parametric ones. Gianola (2006) used simulations and concluded that non-parametric RKHS

model outperformed the parametric standard additive genetic model for additive by additive gene action. Crossa et al. (2010) reported that the RKHS models outperformed the BLUP. Crossa et al. (2013) compared GBLUP with the RKHS and concluded that there was no clear superiority of either of the models although the RKHS-M performed slightly better than the GBLUP. Howard et al. (2014) also reported that the non-parametric models performed well when the underlying genetic architecture was entirely based on epistasis. But for the traits that we analysed in this study, either a negligible effect of epistasis or the equivalence of the RKHS-M to the GBLUP when the kernel used in RKHS is a Gaussian kernel (K=G) (Jiang and Reif, 2015), led to similar accuracies.

We also observed that the RKHS-P performed well and the increase in accuracies using genome-wide marker based models was only in 4.44% (ranged between -20.94 and 38.2%). But the general expectation is that, the pedigree based relationship would predict a 50% relationship between full-sibs and 25% relationship between half-sibs, while the genomic-based relationship would predict the allele sharing (within family variation) with better accuracy (Hayes and Goddard, 2010). This is because it exploits the Mendelian sampling term that occurs during the formation of gametes and captures the realized relationship matrix instead of the average relationship matrix obtained from the pedigree (Daetwyler et al. 2007; Goddard and Hayes 2007; Hayes et al. 2009; Villanueva et al. 2005). Crossa et al. (2010) reported that the gain in using markers compared to the pedigree was 7.7 to 35.7%. Wolc et al. (2011a) showed that marker estimated BVs were more persistent over generations compared to the pedigree estimated BVs in layer chickens. In another study, Wolc et al. (2011b) also reported that marker based methods had higher accuracies than the pedigree based method. Spindel et al. (2015) reported that GS models were superior to the pedigree based prediction in rice for yield, height and flowering time. The benefits of using the G-matrix are manifold: (i) The G-matrix can differentiate sibs and can help

avoid selecting closely related sibs together (Daetwyler et al., 2007) (ii) The G-matrix can provide some prediction accuracies compared to the pedigree (almost zero) when distant/unrelated individuals are involved (van der Werf, 2009) (iii) The G-matrix can perform better when the pedigree is shallow (goes back to only a few generations) (iv) The G-matrix can correct for pedigree errors (Munoz et al., 2014). Nevertheless, the fact that genotypes can also contain errors cannot be overlooked. We attribute the high accuracies obtained with the pedigree in our study to several reasons: (i) CIMMYT maintains an excellent pedigree recording system that goes back several generations. (ii) The family sizes were small and except for large family sizes (with considerable Mendelian segregation), the advantage of using markers over the pedigree is expected to be small. (iii) Dense marker coverage is essential to maximize the number of QTL that will be in LD with at least one marker that in turn is governed by the rate of decay of LD in the genome (Heffner et al. 2009). In this study, the large number of markers would seem to provide excellent genome coverage. But, it is possible that these markers inadequately cover some major regions associated with the trait resulting in lower genomic prediction accuracies. Combining marker and pedigree data can partially correct for gaps in genome coverage. (iv) Another possibility is that, in the highly inbred lines we used, inbreeding resulted in the loss of alleles reducing the Mendelian sampling variance as suggested by Daetwyler et al. (2007). (v) Full-sibs in both the training and validation sets could have lead to higher accuracies with the pedigree, but this might not work as well for lines in early generations.

The RKHS-MP model performed better than just the pedigree and markers alone and gave the highest accuracies for most datasets which is consistent with several studies (Burgueño et al. 2012; Crossa et al. 2010, 2013; de los Campos et al. 2009; Perez et al. 2010). The average increase in accuracy using the RKHS-MP model over RKHS-P was 9.3% (ranged between -1.56 and

32.35%) and over the RKHS-M was 5.23% (ranged between -4.26 and 22.03%). Hence, despite, the pedigree being remarkably robust, it was clear that molecular markers can complement the pedigree to enhance breeding progress. Certain folds were predicted with a higher accuracy using the pedigree and vice versa, although the average accuracies were similar (data not shown). While it would be ideal to use both pedigree and markers to obtain the relationship matrix as suggested by Meuwissen (2007), consideration should be given to how informative the pedigrees are versus the cost of markers to make breeding decisions. However, there is a level of redundancy between the regression on the markers and that on the pedigree, and as a result there might be only a small advantage of considering them together (Habier et al. 2009).

Although the IBWSNs were composed of a set of diverse lines involving several crosses between different parents, the ability to detect significant associations and predict resistance was not high in some datasets, especially where the resistance was quantitative. This was probably due to the lack of variability in these highly selected elite lines that resulted in low power. Hence, the issue is how to effectively implement genomic selection in later generations for traits with limited genetic variability. One strategy that can be applied to a large scale breeding program is to develop a training population of a few hundred carefully chosen diverse fixed lines/varieties that vary widely for resistance to diseases of interest, are closely related to the breeding germplasm, and are grown in a managed nursery. These can be genotyped once and phenotyped for the desired diseases each season at a reasonable cost. Also, new lines from the most recent germplasm can be added to the training population so that prediction models for the highly selected late generation lines will provide more accurate results.

We also observed large differences between the different years/locations/replications for the traits. Several studies have focused on the incorporation of the genotype x environment (G x

E) component in predictions (Burgueño et al. 2012; Heslot et al. 2013a, 2013b; Jarquin et al. 2014; Lopez-Cruz et al. 2015) and it is important to consider the number of environments (years/locations/replications) that should be used for training the model such that it is reasonably stable within and across environments. With whole genome marker genotypes, the unit of replication is the allele and not the genotype *per se*. So, using phenotyping strategies that can maximize the replication of alleles over the replication of individuals (Heslot et al. 2015) is important. In conclusion, our study clearly indicates that for quantitative traits, using genome-wide marker based models maximizes genetic gain using molecular markers compared to marker-assisted selection. GS extends marker-assisted selection to a genome-wide scale and helps to make more accurate and informed breeding decisions for quantitative traits, thus advancing the revolution that molecular markers have brought to crop improvement.

REFERENCES

Abeysekara, N.S., T.L. Friesen, B. Keller, and J.D. Faris. 2009. Identification and characterization of a novel host-toxin interaction in the wheat-*Stagonospora nodorum* pathosystem. Theor Appl Genet 120: 117–126.

Asoro, F.G., M.A. Newell, W.D. Beavis, M.P. Scott, and J.-L. Jannink. 2011. Accuracy and Training Population Design for Genomic Selection on Quantitative Traits in Elite North American Oats. Plant Genome 4: 132–144.

Basnet, B.R., R.P. Singh, S.A. Herrera-Foessel, A.M.H. Ibrahim, J. Huerta‐Espino, V. Calvo-Salazar, et al. 2013. Genetic Analysis of Adult Plant Resistance to Yellow Rust and Leaf Rust in Common Spring Wheat Quaiu 3. Plant Dis 97(6): 728–736.

Bernardo, R. 2014. Genomewide selection when major genes are known. Crop Sci 54: 68–75.

Bernardo, R., and J. Yu. 2007. Prospects for genomewide selection for quantitative traits in maize. Crop Sci 47: 1082–1090.

Bhathal, J.S., R. Loughman, and J. Speijers. 2003. Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and Septoria nodorum blotch. Eur J Plant Pathol 109: 435–443.

Bolton, M.D., J.A. Kolmer, and D.F. Garvin. 2008. Wheat leaf rust caused by Puccinia triticina. Mol Plant Pathol 9: 563–575.

Box, G.E.P., and D.R. Cox. 1964. An Analysis of Transformations. J R Stat Soc Ser B 26(2): 211–252.

Brading, P.A., E.C.P. Verstappen, G.H.J. Kema, and J.K.M. Brown. 2002. A Gene-for-Gene Relationship Between Wheat and *Mycosphaerella graminicola*, the Septoria Tritici Blotch Pathogen. Phytopathology 92(4): 439–445.

Brown, J.K.M., L. Chartrain, P. Lasserre-Zuber, and C. Saintenac. 2015. Genetics of resistance to *Zymoseptoria tritici* and applications to wheat breeding. Fungal Genet Biol 79: 33–41.

Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa. 2012. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Sci 52: 707–719.

Chapman, J.A., M. Mascher, A. Buluç, K. Barry, E. Georganas, A. Session, et al. 2015. A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. Genome Biol 16(1): 26.

Chen, X.M. 2005. Epidemiology and control of stripe rust [Puccinia striiformis f . sp . tritici] on wheat. Can J Plant Pathol 27: 314–337.

Clark, S.A., J.M. Hickey, and J.H. van der Werf. 2011. Different models of genetic variation and

their effect on genomic evaluation. Genet Sel Evol 43: 18.

Crook, A.D., T.L. Friesen, Z.H. Liu, P.S. Ojiambo, and C. Cowger. 2012. Novel Necrotrophic Effectors from *Stagonospora nodorum* and Corresponding Host Sensitivities in Winter Wheat Germplasm in the Southeastern United States. Phytopathology 102: 498–505.

Crossa, J., Y. Beyene, S. Kassa, P. Perez, J.M. Hickey, C. Chen, et al. 2013. Genomic Prediction in Maize Breeding Populations with Genotyping-by-Sequencing. G3 Genes|Genomes|Genetics 3(11): 1903–1926.

Crossa, J., J. Burgueño, S. Dreisigacker, M. Vargas, S.A. Herrera-Foessel, M. Lillemo, et al. 2007. Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics 177(3): 1889–1913.

Crossa, J., G. de los Campos, P. Pérez, D. Gianola, J. Burgueño, J.L. Araus, et al. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186: 713–724.

Crossa, J., P. Pérez, J. Hickey, J. Burgueño, L. Ornella, J. Cerón-Rojas, et al. 2014. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity (Edinb) 112: 48–60.

Daetwyler, H.D., J.M. Hickey, J.M. Henshall, S. Dominik, B. Gredler, J.H.J. Van Der Werf, et al. 2010a. Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population. Anim Prod Sci 50: 1004–1010.

Daetwyler, H.D., R. Pong-Wong, B. Villanueva, and J.A. Woolliams. 2010b. The impact of genetic architecture on genome-wide evaluation methods. Genetics 185: 1021–1031.

Daetwyler, H.D., B. Villanueva, P. Bijma, and J.A. Woolliams. 2007. Inbreeding in genome-wide selection. J Anim Breed Genet 124: 369–376.

Dekkers, J.C.M., and F. Hospital. 2002. The Use of Molecular Genetics in the Improvement of

Agricultural Populations. Nat Rev Genet 3(1): 22–32.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, et al. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS One 6(5): e19379.

Endelman, J.B. 2011. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. Plant Genome 4: 250–255.

Eyal, Z., A.L. Scharen, J.M. Prescott, and M. van Ginkel. 1987. The Septoria Diseases of Wheat: Concepts and methods of disease management. CIMMYT., Mexico, D.F.

Faris, J.D., Z. Liu, and S.S. Xu. 2013. Genetics of tan spot resistance in wheat. Theor Appl Genet 126: 2197–2217.

Feng, J., H. Ma, and G.R. Hughes. 2004. Genetics of resistance to Stagonospora nodorum blotch of hexaploid wheat. Crop Sci 44: 2043–2048.

Fisher, R.A. 1918. The Correlation between Relatives on the Supposition of Mendelian Inheritance. Trans R Soc Edinburgh 52: 399–433.

Flor, H.H. 1956. The complementary genic systems in flax and flax rust. Adv Genet 8: 29–54.

Fones, H., and S. Gurr. 2015. The impact of Septoria tritici Blotch disease on wheat: An EU perspective. Fungal Genet Biol 79: 3–7.

Friesen, T.L., C. Chu, S.S. Xu, and J.D. Faris. 2012. SnTox5-*Snn5*: A novel *Stagonospora nodorum* effector-wheat gene interaction and its relationship with the SnToxA-*Tsn1* and SnTox3-*Snn3-B1* interactions. Mol Plant Pathol 13: 1101–1109.

Friesen, T.L., S.W. Meinhardt, and J.D. Faris. 2007. The *Stagonospora nodorum*-wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. Plant J 51: 681–692.

Friesen, T.L., E.H. Stukenbrock, Z. Liu, S. Meinhardt, H. Ling, J.D. Faris, et al. 2006. Emergence of a new disease as a result of interspecific virulence gene transfer. Nat Genet 38(8): 953–956.

Friesen, T.L., Z. Zhang, P.S. Solomon, R.P. Oliver, and J.D. Faris. 2008. Characterization of the interaction of a novel *Stagonospora nodorum* host-selective toxin with a wheat susceptibility gene. Plant Physiol 146: 682–693.

Gao, Y., J.D. Faris, Z. Liu, Y.M. Kim, R.A. Syme, R.P. Oliver, et al. 2015. Identification and Characterization of the SnTox6-Snn6 Interaction in the Parastagonospora nodorum – Wheat Pathosystem. Mol Plant-Microbe Interact 28(5): 615–625.

Ghaffary, S.M.T., J.D. Faris, T.L. Friesen, R.G.F. Visser, T.A.J. van der Lee, O. Robert, et al. 2012. New broad-spectrum resistance to septoria tritici blotch derived from synthetic hexaploid wheat. Theor Appl Genet 124: 125–142.

Gianola, D. 2013. Priors in whole-genome regression: The Bayesian alphabet returns. Genetics 194(3): 573–596.

Gianola, D., R.L. Fernando, and A. Stella. 2006. Genomic-Assisted Prediction of Genetic Value With Semiparametric Procedures. Genetics 173: 1761–1776.

Gianola, D., and J.B.C.H.M. van Kaam. 2008. Reproducing Kernel Hilbert Spaces Regression Methods for Genomic Assisted Prediction of Quantitative Traits. Genetics 178(4): 2289–2303.

Gilchrist-Saavedra, L., G. Fuentes-Dávila, C. Martínez-Cano, R.M. López-Atilano, E. Duveiller, R.P. Singh, et al. 2006. Practical guide to the identification of selected diseases of wheat and barley. Second. CIMMYT, Mexico, D.F.

Gilmour, A.R., R. Thompson, and B.R. Cullis. 1995. Average Information REML: An Efficient

Algorithm for Variance Parameter Estimation in Linear Mixed Models. Biometrics 51(4): 1440–1450.

van Ginkel, M., and S. Rajaram. 1993. Breeding for durable resistance to diseases in wheat an additional perspective. Durab Dis Resist: 259–272.

Goddard, M. 2009. Genomic selection: prediction of accuracy and maximisation of long term response. Genetica 136(2): 245–257.

Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. J Anim Breed Genet 124(6): 323–330.

Goodwin, S.B. 2007. Back to basics and beyond: Increasing the level of resistance to Septoria tritici blotch in wheat. Australas Plant Pathol 36: 532–538.

Habier, D., R.L. Fernando, and J.C.M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177(4): 2389–2397.

Habier, D., R.L. Fernando, and J.C.M. Dekkers. 2009. Genomic Selection Using Low-Density Marker Panels. Genetics 182(1): 343–353.

Habier, D., R.L. Fernando, and D.J. Garrick. 2013. Genomic BLUP decoded: A look into the black box of genomic prediction. Genetics 194(3): 597–607.

Habier, D., R.L. Fernando, K. Kizilkaya, and D.J. Garrick. 2011. Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics 12: 186.

Hayes, B., and M. Goddard. 2010. Genome-wide association and genomic selection in animal breeding. Genome 53(11): 876–83.

Hayes, B.J., P.M. Visscher, and M.E. Goddard. 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genet Res (Camb) 91(1): 47–60.

Heffner, E.L., J.L. Jannink, H. Iwata, E. Souza, and M.E. Sorrells. 2011a. Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci 51: 2597–2606.

Heffner, E.L., J. Jannink, and M.E. Sorrells. 2011b. Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. Plant Genome 4(1): 65–75.

Heffner, E.L., M.E. Sorrells, and J.-L. Jannink. 2009. Genomic Selection for Crop Improvement. Crop Sci 49: 1–12.

Heslot, N., D. Akdemir, M.E. Sorrells, and J.L. Jannink. 2013a. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor Appl Genet: 1–18.

Heslot, N., J.L. Jannink, and M.E. Sorrells. 2013b. Using genomic prediction to characterize environments and optimize prediction accuracy in applied breeding data. Crop Sci 53(June): 921–933.

Heslot, N., J.-L. Jannink, and M.E. Sorrells. 2015. Perspectives for Genomic Selection Applications and Research in Plants. Crop Sci 55(february): 1–12.

Heslot, N., J. Rutkoski, J. Poland, J.L. Jannink, and M.E. Sorrells. 2013c. Impact of Marker Ascertainment Bias on Genomic Selection Accuracy and Estimates of Genetic Diversity. PLoS One 8(9): e74612.

Heslot, N., H.-P. Yang, M.E. Sorrells, and J.-L. Jannink. 2012. Genomic Selection in Plant Breeding: A Comparison of Models. Crop Sci 52: 146–160.

Hill, W.G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor Appl Genet 38(6): 226–231.

Hoerl, A.E., and R.W. Kennard. 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12(1): 55–67.

Howard, R., A.L. Carriquiry, and W.D. Beavis. 2014. Parametric and Nonparametric Statistical Methods for Genomic Selection of Traits with Additive and Epistatic Genetic Architectures.

G3 Genes|Genomes|Genetics 4(6): 1027–1046.

Huerta-Espino, J., R.P. Singh, S. Germán, B.D. McCallum, R.F. Park, W.Q. Chen, et al. 2011. Global status of wheat leaf rust caused by *Puccinia triticina*. Euphytica 179(1): 143–160.

Jiang, Y., and J.C. Reif. 2015. Modelling Epistasis in Genomic Selection. Genetics: genetics.115.177907-.

Johnson, R. 1984. A critical analysis of durable resistance. Annu Rev Phytopathol 22: 309–330.

Kizilkaya, K., R.L. Fernando, and D.J. Garrick. 2010. Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. J Anim Sci 88: 544–551.

Kolmer, J.A., J.A. Anderson, and J.M. Flor. 2010. Chromosome location, linkage with simple sequence repeat markers, and leaf rust resistance conditioned by gene *Lr63* in wheat. Crop Sci 50(6): 2392–2395.

Kruijer, W., M.P. Boer, M. Malosetti, P.J. Flood, B. Engel, R. Kooke, et al. 2015. Marker-Based Estimation of Heritability in Immortal Populations. Genetics 199(2): 379–398.

Lamari, L., and C.C. Bernier. 1989a. Evaluation of wheat lines and cultivars to tan spot [*Pyrenophora tritici-repentis*] based on lesion type. Can J Plant Pathol 11: 49–56.

Lamari, L., and C.C. Bernier. 1989b. Virulence of isolates of *Pyrenophora tritici-repentis* on 11 wheat cultivars and cytology of the differential host reactions. Can J Plant Pathol 11(3): 284–290.

Lan, C.X., R.P. Singh, J. Huerta-Espino, V. Calvo-Salazar, and S.A. Herrera-Foessel. 2014. Genetic Analysis of Resistance to Leaf Rust and Stripe Rust in Wheat Cultivar Francolin#1. Plant Dis 98(9): 1227–1234.

Leonard, K.J., and L.J. Szabo. 2005. Stem rust of small grains and grasses caused by Puccinia

graminis. Mol Plant Pathol 6(2): 99–111.

Li, H., P. Vikram, R.P. Singh, A. Kilian, J. Carling, J. Song, et al. 2015. A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. BMC Genomics 16: 216.

Liu, Z.H., J.D. Faris, S.W. Meinhardt, S. Ali, J.B. Rasmussen, and T.L. Friesen. 2004. Genetic and Physical Mapping of a Gene Conditioning Sensitivity in Wheat to a Partially Purified Host-Selective Toxin Produced by *Stagonospora nodorum*. Phytopathology 94(5): 1056–1060.

Liu, Z., T.L. Friesen, H. Ling, S.W. Meinhardt, R.P. Oliver, J.B. Rasmussen, et al. 2006. The *Tsn1*-ToxA interaction in the wheat-*Stagonospora nodorum* pathosystem parallels that of the wheat-tan spot system. Genome 49: 1265–1273.

Liu, Z., Z. Zhang, J.D. Faris, R.P. Oliver, R. Syme, M.C. McDonald, et al. 2012. The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring *Snn1*. PLoS Pathog 8(1): e1002467.

Lopez-Cruz, M., J. Crossa, D. Bonnett, S. Dreisigacker, J. Poland, J.-L. Jannink, et al. 2015. Increased Prediction Accuracy in Wheat Breeding Trials Using a Marker x Environment Interaction Genomic Selection Model. G3 Genes|Genomes|Genetics 5(4): 569–82.

Lorenz, A.J., S. Chao, F.G. Asoro, E.L. Heffner, T. Hayashi, H. Iwata, et al. 2011. Genomic Selection in Plant Breeding: Knowledge and prospects. Adv Agron 110: 77–123.

Lorenzana, R.E., and R. Bernardo. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. Theor Appl Genet 120(1): 151–161.

Lorenz, A.J., K.P. Smith, and J.L. Jannink. 2012. Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. Crop Sci 52: 1609–1621.

De los Campos, G., D. Gianola, G.J.M. Rosa, K. a Weigel, and J. Crossa. 2010. Semi-parametric

genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. Genet Res (Camb) 92(4): 295–308.

de los Campos, G., J.M. Hickey, R. Pong-Wong, H.D. Daetwyler, and M.P.L. Calus. 2013a. Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. Genetics 193: 327–345.

de los Campos, G., H. Naya, D. Gianola, J. Crossa, A. Legarra, E. Manfredi, et al. 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics 182: 375–385.

de los Campos, G., A.I. Vazquez, R. Fernando, Y.C. Klimentidis, and D. Sorensen. 2013b. Prediction of complex human traits using the genomic best linear unbiased predictor. PLoS Genet 9(7): e1003608.

Luan, T., J.A. Woolliams, S. Lien, M. Kent, M. Svendsen, and T.H.E. Meuwissen. 2009. The accuracy of genomic selection in Norwegian red cattle assessed by cross-validation. Genetics 183: 1119–1126.

Marasas, C.N., M. Smale, and R.P. Singh. 2004. The Economic Impact in Developing Countries of Leaf Rust Resistance Breeding in CIMMYT-Related spring bread wheat. CIMMYT Mex DF.

McIntosh, R.A., J. Dubcovsky, W. Rogers, C. Morris, R. Appels, and X.C. Xia. 2016. Catalogue of gene symbols for wheat: 2015-2016 supplement.

McIntosh, R., C. Wellings, and R. Park. 1995. Wheat rusts: an atlas of resistance genes. CSIRO Publishing.

McNeal, F.H., C.F. Konzak, E.P. Smith, W.S. Tate, and T.S. Russell. 1971. A uniform system for recording and processing cereal research data. Agricultural Research Service, United States

Department of Agriculture, [Beltsville Md.].

Meien-Vogeler, F., G. Bartels, and H. Fehrmann. 1994. Distribution of *S. tritici* on wheat in Germany and its regional importance. p. 299–300. *In* Proceedings of the 4th International Septoria of Cereals Workshop. IHAR, Radzikow, Poland.

Meuwissen, T. 2007. Genomic selection : marker assisted selection on a genome wide scale. J Anim Breed Genet 124: 321–322.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829.

Moser, G., B. Tier, R.E. Crump, M.S. Khatkar, and H.W. Raadsma. 2009. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. Genet Sel Evol 41: 56.

Muir, W.M. 2007. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. J Anim Breed Genet 124: 342–355.

Munoz, P.R., M.F.R. Resende, D.A. Huber, T. Quesada, M.D. V. Resende, D.B. Neale, et al. 2014. Genomic Relationship Matrix for Correcting Pedigree Errors in Breeding Populations: Impact on Genetic Parameters and Genomic Selection Accuracy. Crop Sci 53: 1115–1123.

Murray, G.M., and J.P. Brennan. 2009. Estimating disease losses to the Australian wheat industry. Australas Plant Pathol 38(6): 558–570.

Newcomb, M., P.D. Olivera, M.N. Rouse, L.J. Szabo, J. Johnson, S. Gale, et al. 2016. Kenyan Isolates of Puccinia graminis f. sp. tritici from 2008 to 2014 : Virulence to SrTmp in the Ug99 Race Group and Implications for Breeding Programs. Phytopathology 106(7): 729–736.

O'Driscoll, A., S. Kildea, F. Doohan, J. Spink, and E. Mullins. 2014. The wheat-Septoria conflict:

A new front opening up? Trends Plant Sci 19(9): 602–610.

Ornella, L., S. Singh, P. Perez, J. Burgueño, R. Singh, E. Tapia, et al. 2012. Genomic Prediction of Genetic Values for Resistance to Wheat Rusts. Plant Genome 5: 136–148.

Orton, E.S., S. Deller, and J.K.M. Brown. 2011. *Mycosphaerella graminicola*: From genomics to disease control. Mol Plant Pathol 12(5): 413–424.

Park, T., and G. Casella. 2008. The Bayesian Lasso. J Am Stat Assoc 103(482): 681–686.

Perez-Rodriguez, P., D. Gianola, J.M. Gonzalez-Camacho, J. Crossa, Y. Manes, and S. Dreisigacker. 2013. Comparison Between Linear and Non-parametric Regression Models for Genome-Enabled Prediction in Wheat. G3 Genes|Genomes|Genetics 2(12): 1595–1605.

Pérez, P., and G. de Los Campos. 2014. Genome-Wide Regression and Prediction with the BGLR Statistical Package. Genetics 198: 483–495.

Perez, P., G. de los Campos, J. Crossa, and D. Gianola. 2010. Genomic-Enabled Prediction Based on Molecular Markers and Pedigree Using the Bayesian Linear Regression Package in R. Plant Genome 3: 106–116.

Peterson, R.F., A.B. Campbell, and A.E. Hannah. 1948. A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. Can J Res 26c(5): 496–500.

Piepho, H.P. 2009. Ridge Regression and Extensions for Genomewide Selection in Maize. Crop Sci 49: 1165–1176.

Piyasatian, N., R.L. Fernando, and J.C.M. Dekkers. 2007. Genomic selection for marker-assisted improvement in line crosses. Theor Appl Genet 115: 665–674.

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7(2): e32253.

Polley, R.W., and M.R. Thomas. 1991. Surveys of diseases of winter wheat in England and Wales, 1976–1988. Ann Appl Biol 119(1): 1–20.

Resende, M.F.R., P. Munoz, M.D. V. Resende, D.J. Garrick, R.L. Fernando, J.M. Davis, et al. 2012. Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (*Pinus taeda* L.). Genetics 190: 1503–1510.

Roelfs, A.P., R.P. Singh, and E.E. Saari. 1992. Rust Diseases of Wheat: Concepts and methods of disease management. CIMMYT, Mexico, D.F.

de Roos, A.P.W., B.J. Hayes, and M.E. Goddard. 2009. Reliability of Genomic Predictions Across Multiple Populations. Genetics 183: 1545–1553.

Rouse, M.N., L.E. Talbert, D. Singh, and J.D. Sherman. 2014. Complementary epistasis involving Sr12 explains adult plant resistance to stem rust in Thatcher wheat (Triticum aestivum L.). Theor Appl Genet 127(7): 1549–1559.

Rutkoski, J., J. Benson, Y. Jia, G. Brown-Guedira, J.-L. Jannink, and M. Sorrells. 2012. Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. Plant Genome J 5(2): 51.

Rutkoski, J.E., E.L. Heffner, and M.E. Sorrells. 2011. Genomic selection for durable stem rust resistance in wheat. Euphytica 179: 161–173.

Rutkoski, J.E., J.A. Poland, R.P. Singh, J. Huerta-Espino, S. Bhavani, H. Barbier, et al. 2014. Genomic Selection for Quantitative Adult Plant Stem Rust Resistance in Wheat. Plant Genome 7(3).

Saari, E.E., and J. Prescott. 1975. A scale for appraising the foliar intensity of wheat diseases. Plant Dis Report 59(5): 376–381.

Sansaloni, C., C. Petroli, D. Jaccoud, J. Carling, F. Detering, D. Grattapaglia, et al. 2011. Diversity

Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. BMC Proc 5: P54.

Shabeer, A., and W.W. Bockus. 1988. Tan spot effects on yield and yield components relative to growth stage in winter wheat. Plant Dis 72: 599–602.

Shi, G., T.L. Friesen, J. Saini, S.S. Xu, J.B. Rasmussen, and J.D. Faris. 2015. The Wheat Gene *Snn7* Confers Susceptibility on Recognition of the *Parastagonospora nodorum* Necrotrophic Effector SnTox7. Plant Genome 8(2).

Singh, P.K., E. Duveiller, and R.P. Singh. 2011. Evaluation of CIMMYT germplasm for resistance to leaf spotting diseases of wheat. Czech J Genet Plant Breed 47: S102–S108.

Singh, R.P., D.P. Hodson, Y. Jin, E.S. Lagudah, M.A. Ayliffe, S. Bhavani, et al. 2015. Emergence and Spread of New Races of Wheat Stem Rust Fungus: Continued Threat to Food Security and Prospects of Genetic Control. Phytopathology: PHYTO01150030FI.

Singh, P.K., M. Mergoum, T.B. Adhikari, S.F. Kianian, and E.M. Elias. 2006. Chromosomal location of genes for seedling resistance to tan spot and Stagonospora nodorum blotch in tetraploid wheat. Euphytica 155: 27–34.

Spindel, J., H. Begum, D. Akdemir, P. Virk, B. Collard, E. Redona, et al. 2015. Genomic Selection and Association Mapping in Rice (*Oryza sativa*): Effect of Trait Genetic Architecture, Training Population Composition, Marker Number and Statistical Model on Accuracy of Rice Genomic Selection in Elite, Tropical Rice Breeding Line. PLoS Genet 11(2): e1004982.

Tibshirani, R. 1996. Regression Selection and Shrinkage via the Lasso. J R Stat Soc B 58: 267–288.

Toosi, A., R.L. Fernando, and J.C.M. Dekkers. 2010. Genomic selection in admixed and crossbred

populations. J Anim Sci 88(1): 32–46.

Torriani, S.F.F., J.P.E. Melichar, C. Mills, N. Pain, H. Sierotzki, and M. Courbot. 2015. *Zymoseptoria tritici*: A major threat to wheat production, integrated approaches to control. Fungal Genet Biol 79: 8–12.

Vanderplank, J.E. 1963. Plant Diseases: Epidemics and Control. Academic Press, NY.

VanRaden, P.M. 2008. Efficient Methods to Compute Genomic Predictions. J Dairy Sci 91(11): 4414–4423.

Villanueva, B., R. Pong-Wong, J. Fernández, and M.A. Toro. 2005. Benefits from marker-assisted selection under an additive polygenic genetic model. J Anim Sci 83: 1747–1752.

van der Werf, J.H.J. 2009. Potential benefit of genomic selection in sheep. Proc Assoc Adv Anim Breed Genet 18: 38–41.

Whittaker, J.C., R. Thompson, and M.C. Denham. 2000. Marker-assisted selection using ridge regression. Genet Res 75: 249–252.

Wolc, A., J. Arango, P. Settar, J.E. Fulton, N.P. O'Sullivan, R. Preisinger, et al. 2011a. Persistence of accuracy of genomic estimated breeding values over generations in layer chickens. Genet Sel Evol 43(23).

Wolc, A., C. Stricker, J. Arango, P. Settar, J.E. Fulton, N.P. O'Sullivan, et al. 2011b. Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. Genet Sel Evol 43(1): 5.

De Wolf, E.D., R.J. Effertz, S. Ali, and L.J. Francl. 1998. Vistas of tan spot research. Can J Plant Pathol 20(4): 349–370.

Wong, C.K., and R. Bernardo. 2008. Genomewide selection in oil palm: Increasing selection gain per unit time and cost with small populations. Theor Appl Genet 116(6): 815–824.

Xu, S. 2003. Estimating polygenic effects using markers of the entire genome. Genetics 163: 789–801.

Yang, J., B. Benyamin, B.P. McEvoy, S. Gordon, A.K. Henders, D.R. Nyholt, et al. 2010. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42(7): 565–569.

Yu, L.X., A. Lorenz, J. Rutkoski, R.P. Singh, S. Bhavani, J. Huerta-Espino, et al. 2011. Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. Theor Appl Genet 123: 1257–1268.

Zadoks, J.C., T.T. Chang, and C.F. Konzak. 1974. A Decimal Code for the Growth Stages of Cereals. Weed Res 14(6): 415–421.

Zhang, Z., T.L. Friesen, S.S. Xu, G. Shi, Z. Liu, J.B. Rasmussen, et al. 2011. Two putatively homoeologous wheat genes mediate recognition of SnTox3 to confer effector-triggered susceptibility to *Stagonospora nodorum*. Plant J 65: 27–38.

Zhao, Y., J. Zeng, R. Fernando, and J.C. Reif. 2013. Genomic Prediction of Hybrid Wheat Performance. Crop Sci 53(3): 802–810.

Zhong, S., J.C.M. Dekkers, R.L. Fernando, and J.-L. Jannink. 2009. Factors Affecting Accuracy From Genomic Selection in Populations Derived From Multiple Inbred Lines: A Barley Case Study. Genetics 182(1): 355–364.

Table S1.1: Basic local alignment search tool (BLAST) results for the significant markers in the 45th and 46th international bread wheat screening nursery (IBWSN)

Marker	Subject[a]	Score	Identities (Query length)	Percentage	Expect
GBS_24751	gbsHWWAMP38350	113	63/64 (64)	98	6.00E-23
	gbsCNLmaster31897	113	63/64 (64)	98	6.00E-23
	2BS_5218802	111	63/64 (64)	98	2.00E-22
GBS_37247	WCSS1_contig470290_1DS-434	113	63/64 (64)	98	6.00E-23
GBS_8842	WCSS1_contig3334901_3AS-4592	113	63/64 (64)	98	6.00E-23
	synopGBS105165	111	62/63 (64)	98	2.00E-22
GBS_30281	gbsHWWAMP47807	113	63/64 (64)	98	6.00E-23
	WCSS1_contig7120458_4AL-2167	113	63/64 (64)	98	6.00E-23
GBS_22856	WCSS1_contig10511286_3B-5360	113	63/64 (64)	98	6.00E-23
GBS_36529	WCSS1_contig10759567_3B-1965	113	63/64 (64)	98	6.00E-23
	synopGBS125790	111	62/63 (64)	98	2.00E-22
GBS_2454	WCSS1_contig2284473_5BS-12686	114	63/63 (64)	100	2.00E-23
	gbsHWWAMP2529	113	63/64 (64)	98	6.00E-23
GBS_13047	gbsHWWAMP18106	113	63/64 (64)	98	6.00E-23
GBS_6432	gbsHWWAMP7358	113	63/64 (64)	98	6.00E-23
	WCSS1_contig5219749_2AS-4945	113	63/64 (64)	98	6.00E-23

GBS_702	gbsHWWAMP627	113	63/64 (64)	98	6.00E-23
	WCSS1_contig5304580_2AS-10182	109	62/64 (64)	97	7.00E-22
GBS_19971	WCSS1_contig1905752_1DS-4628	114	63/63 (64)	100	2.00E-23
GBS_28186	gbsHWWAMP44231	113	63/64 (64)	98	6.00E-23
	gbsCNLmaster36568	113	63/64 (64)	98	6.00E-23
	synopGBS121579	111	62/63 (64)	98	2.00E-22
GBS_28376	WCSS1_contig1898017_1DS-2235	111	63/64 (64)	98	2.00E-22
GBS_40747	gbsHWWAMP55344	113	63/64 (64)	98	6.00E-23
	2DS_5390754	116	64/64 (64)	100	
GBS_18425	gbsCNLmaster24464	116	64/64 (64)	100	5.00E-24
	WCSS1_contig3419689_3AS-1090	116	64/64 (64)	100	5.00E-24
GBS_2400	gbsHWWAMP2465	116	64/64 (64)	100	5.00E-24
	WCSS1_contig3361063_3AS-2705	113	63/64 (64)	98	6e-23
GBS_1491	gbsHWWAMP1393	116	64/64 (64)	100	5.00E-24
GBS_23856	gbsHWWAMP37196	113	63/64 (64)	98	6e-23
GBS_28025	WCSS1_contig3042477_6BS-5453	113	63/64 (64)	98	6e-23
GBS_20060	WCSS1_contig2078323_6DS-22086	113	63/64 (64)	98	6e-23

[a] Marker prefixes and the population they were genotyped in or the map they are available in: gbsHWWAMP is for markers from the hard winter wheat association mapping panel; gbsCNLmaster is for markers from Cornell wheat master nursery; synopGBS is for markers in the Synthetic and Opata map (Poland et al., 2012) and WCSS1_contig is for markers in the CSS GBS 2014 physical map where 'WCSS1' stands for wheat chromosome survey sequence.

CHAPTER 2

COMPARISON OF MODELS AND WHOLE-GENOME PROFILING APPROACHES FOR GENOMIC-ENABLED PREDICTION OF SEPTORIA TRITICI BLOTCH, STAGONOSPORA NODORUM BLOTCH AND TAN SPOT RESISTANCE IN WHEAT

ABSTRACT

The leaf spotting diseases in wheat that include Septoria tritici blotch (STB) caused by *Zymoseptoria tritici*, Stagonospora nodorum blotch (SNB) caused by *Parastagonospora nodorum* and tan spot (TS) caused by *Pyrenophora tritici-repentis* pose challenges to breeding programs in selecting for resistance. A promising approach that could enable selection prior to phenotyping is genomic selection that uses genome-wide markers to estimate breeding values for quantitative traits. To evaluate this approach for seedling and/or adult plant resistance (APR) to STB, SNB and TS, we compared the predictive ability of least-squares (LS) approach with genomic-enabled prediction models including genomic BLUP, Bayesian ridge regression, Bayes A, Bayes B, Bayes Cπ, Bayesian least absolute shrinkage and selection operator and reproducing kernel Hilbert spaces (RKHS) markers, a pedigree-based model (RKHS-pedigree) and RKHS-markers and pedigree (RKHS-MP). We observed that LS gave the lowest prediction accuracies and RKHS-MP, the highest. The genomic-enabled prediction models and RKHS-pedigree gave similar accuracies. The increase in accuracy using genomic prediction models over LS was 48%. The mean genomic prediction accuracies were 0.45 for STB (APR), 0.55 for SNB (seedling), 0.66 for TS (seedling) and 0.48 for TS (APR). We also compared markers from two whole-genome profiling approaches: genotyping by sequencing (GBS) and diversity arrays technology sequencing (DArTseq) for

prediction. While, GBS markers performed slightly better than DArTseq, combining markers from the two approaches did not improve accuracies. We conclude that implementing GS in breeding for these diseases would help to achieve higher accuracies and rapid gains from selection.

ABBREVIATIONS

APR, adult plant response; BA, Bayes A; BB, Bayes B; BC, Bayes Cπ; BL, Bayesian least absolute shrinkage and selection operator; BLUP, best linear unbiased predictor; BRR, Bayesian ridge regression; BV, breeding value; CIMMYT, Centro Internacional de Mejoramiento de Maíz y Trigo; DArTseq, diversity arrays technology sequencing; GBLUP, genomic best linear unbiased predictor; GBS, genotyping by sequencing; IBWSN, international bread wheat screening nursery; IID, independent and identically distributed; LD, linkage disequilibrium; LS, least-squares; NE, necrotrophic effectors; PDA, potato dextrose agar; QTL, quantitative trait loci; rAUDPC, relative area under the disease progression curve; RKHS-M, reproducing kernel Hilbert spaces markers; RKHS-MP, reproducing kernel Hilbert spaces markers and pedigree; RKHS-P, reproducing kernel Hilbert spaces pedigree; RR-BLUP, ridge regression-best linear unbiased prediction; SNB, Stagonospora nodorum blotch; STB, Septoria tritici blotch; TS, tan spot.

INTRODUCTION

The major leaf spotting diseases threatening wheat (*Triticum aestivum* L.) are Septoria tritici blotch (STB) caused by *Zymoseptoria tritici* (Desm.) Quaedvlieg & Crous, Stagonospora nodorum blotch (SNB) caused by *Parastagonospora nodorum* (Berk.) Quaedvlieg, Verkley & Crous and tan spot (TS) caused by *Pyrenophora tritici-repentis* (Died.) Drechsler. Among these,

STB is an important disease in the temperate regions of the world and is considered to be the most damaging disease of wheat in Europe (Eyal et al., 1987; Goodwin, 2007; Orton et al., 2011; O'Driscoll et al., 2014). While the average annual yield loss in the UK was 20% when susceptible lines were not treated with fungicides, only 5-10% loss resulted from using resistant varieties and fungicide treatment (Fones and Gurr, 2015). About 70% ($1.2 billion) of the annual cereal fungicides in the European Union is used for STB management and fungicide resistance in *Z. tritici* populations is widespread (Torriani et al., 2015). This has made genetic resistance the preferred STB management strategy, which can be either qualitative (controlled by large effect major genes that follow the gene-for-gene model) or quantitative (controlled by few to many genes of moderate to small effects) (Brown et al., 2015). Several genes for STB resistance have been reported which include *Stb1-Stb15*, *StbSm3, Stb16q, Stb17, Stb18, StbWW* and *TmStb1* (Brown et al., 2015). Among these, *Stb6* interaction shows a typical gene-for-gene relationship (Brading et al., 2002) and *Stb17* is a gene for quantitative resistance expressed at the adult plant stage (Ghaffary et al., 2012).

Stagonospora nodorum blotch or glume blotch, is an important disease in the warm and moist growing areas of the world that can cause yield losses of up to 31% under high inoculum pressure (Bhathal et al., 2003). The relative importance of the causal necrotroph, *P. nodorum* varies in different parts of the world. It is a major pathogen of winter wheat in the United States (Crook et al., 2012), the second most economically important pathogen in the Western region in Australia (Murray and Brennan, 2009) and there was a shift in its prevalence in Europe when it was overtaken by *Z. tritici* populations in both UK and Germany (Polley and Thomas, 1991; Meien-Vogeler et al., 1994). Tan spot or yellow spot, is another devastating foliar disease that is a serious

constraint to wheat production in Western Australia (Murray and Brennan, 2009). It can result in an average yield loss of 5-10%, but losses up to 50% can occur under conditions favorable for disease development (Shabeer and Bockus, 1988; Lamari and Bernier, 1989a; b; De Wolf et al., 1998). While fungicides and agronomic practices are available for SNB and TS management, the deployment of resistant cultivars is the most sustainable, cost-effective and environment friendly strategy. Host-pathogen interactions for both *P. nodorum* and *P. tritici-repentis,* follow the inverse gene-for-gene model. This involves the recognition of host-specific toxins or necrotrophic effectors (NE) by a host sensitivity gene resulting in a compatible interaction, leading to susceptibility. The non-recognition of the toxin by the host results in an incompatible interaction leading to resistance (Faris et al. 2013). For SNB, several interactions between NE and the host genes have been identified which include, *Snn1* (Liu et al., 2004), *Tsn1* (Friesen et al., 2006; Liu et al., 2006), *Snn2* (Friesen et al., 2007), *Snn3* (Friesen et al., 2008), *Snn3-B1* and *Snn3-D1* (Zhang et al., 2011), *Snn4* (Abeysekara et al., 2009), *Snn5* (Friesen et al., 2012), *Snn6* (Gao et al., 2015) and *Snn7* (Shi et al., 2015). For TS, six qualitative genes, *Tsr1-Tsr6* that interact with a range of host-specific toxins, including ToxA, ToxB and ToxC have been reported (Faris et al., 1996; Singh et al., 2006; Tadesse et al., 2006a and b; Singh et al., 2008; Friesen and Faris, 2004).

Breeding for resistance to wheat leaf spotting diseases is a challenge because of the difficulties in phenotyping, the ephemeral nature of some of the known resistance genes, the emergence of new isolates and the complex inheritance of genetic resistance. Hence, it is important to devise strategies to accelerate breeding for quantitative resistance which is likely to be more durable. One promising approach that could help achieve this is genomic selection (GS) (Meuwissen et al. 2001) which uses dense genome-wide markers to obtain the genomic estimated

BVs of individuals. This enables selection prior to phenotyping, thereby leading to greater rates of genetic gain. The potential of GS to improve quantitative traits in wheat has been demonstrated in many empirical studies (Crossa et al., 2010, 2014; Heslot et al., 2012; Ornella et al., 2012; Rutkoski et al., 2014). In GS, a 'training population' comprising individuals that have been genotyped and phenotyped for traits of interest is used to train a model that is used to predict the BVs of individuals in a 'selection population' that is not phenotyped. While some studies comparing prediction model accuracies have been reported (Lorenzana and Bernardo, 2009; Crossa et al., 2010; Heffner et al., 2011; Heslot et al., 2012), our objective was to compare the predictive ability of the least-squares (LS) approach (where selected loci were used as fixed effects) with genomic-enabled prediction models for STB, SNB and TS resistance. The genomic prediction models evaluated include genomic BLUP (GBLUP), Bayesian ridge regression (BRR), Bayes A (BA), Bayes B (BB), Bayes Cπ (BC), Bayesian least absolute shrinkage and selection operator (BL) and reproducing kernel Hilbert spaces (RKHS) markers (RKHS-M). In addition, we also evaluated a pedigree based model, RKHS pedigree (RKHS-P) and RKHS markers and pedigree (RKHS-MP) that included both the pedigree and marker based relationship matrices. We also compared markers obtained from two whole-genome profiling approaches for genomic prediction: the genotyping by sequencing (GBS) method (Poland et al., 2012) and the diversity arrays technology sequencing (DArTseq), used by diversity arrays technology, Canberra, Australia (http://www.diversityarrays.com/dart-application-dartseq).

MATERIALS AND METHODS

Plant materials

For this study, we used CIMMYT's (Centro Internacional de Mejoramiento de Maíz y Trigo) 45th and 46th international bread wheat screening nurseries (IBWSN) comprising 333 and 313 lines, respectively. The IBWSNs are large screening nurseries that are evaluated in multiple trials in Mexico and cooperating locations globally. They consist of 200 to 400 advanced lines from CIMMYT's bread wheat breeding program. They are expected to have several novel genes for resistance and considerable variation in their BVs, making them ideal for building prediction models.

Disease evaluation and phenotypic data

Adult plant resistance evaluation for Septoria tritici blotch

Adult plant resistance to STB was evaluated at CIMMYT's research station, Toluca, Mexico during the 2011, 2013 and 2014 crop seasons. The inoculum for STB was prepared according to Gilchrist-Saavedra et al. (2006) using a mixture of six aggressive strains: St1 (B1), St2 (P8), St5 (OT), St6 (KK), 64 (St 81.1) and 86 (St 133.4) at a concentration of 1 x 10^7 spores/ml. The nurseries were inoculated 45 days after planting using an ultra-low volume applicator. Two additional applications were made at weekly intervals. A border row of a susceptible spreader variety, Huirivis, and a resistant variety, Murga, was planted surrounding the field. The plants were evaluated using the double-digit scale (00-99) that is a modification of the Saari-Prescott 0-9 scale for rating foliar diseases (Saari and Prescott, 1975). The first digit gives the relative height of the disease spread vertically using the original 0-9 Saari-Prescott scale and the second digit represents the percentage disease severity in terms of 0-9 (Eyal et al., 1987). Three to four evaluations were carried out. The disease severity percentages were calculated from the scores using the formula:

(first score/9) × (second score/9) × 100 and were used to obtain the relative area under the disease progression curve (rAUDPC). In 2014, there was high incidence of stripe rust and the two diseases became nearly inseparable. So, we included stripe rust severity as a covariate in all the models for this year.

Seedling evaluation for Stagonospora nodorum blotch

Seedling resistance to SNB was evaluated in CIMMYT's greenhouses, El Batan, Mexico in 2014. Inoculum production and inoculation were done as described in Singh et al. (2006). The *P. nodorum* isolate *Sn*4 at a concentration of 1 x 10^6 spores/ml was used. Each entry was represented by four seedlings planted in six replications and the check varieties Erik, Glenlea, 6B-662 and 6B-365 were planted every 20 rows. The second leaf of each seedling was scored for SNB disease reaction seven days post inoculation, using the 1 to 5 lesion rating scale (Feng et al. 2004).

Seedling and adult plant resistance evaluation for tan spot

Seedling resistance and APR to TS were evaluated at CIMMYT's greenhouses and fields at El Batan, Mexico, 2014. Race 1 (isolate *Ptr*1) was used and the inoculum was produced by the method described by Singh et al. (2011). The concentration of the inoculum (for both seedling and field inoculation) was adjusted to 4000 conidia/ml. Seedling inoculation and checks were similar to that for *P. nodorum* and were planted in six replications. Seven days post inoculation, the seedlings were rated for disease response based on a 1 to 5 lesion rating scale developed by Lamari and Bernier (1989a). Field inoculation and evaluation was similar to that for STB. A continuous border row of the susceptible spreader, Glenlea, and the resistant variety, Erik was planted

surrounding the field. The double-digit scale was used and the rAUDPC was calculated from four evaluations done at weekly intervals.

The phenotypic distributions for all the diseases were transformed using the boxcox function in the 'R' statistical program.

Genotyping

The two nurseries were genotyped using the GBS method described by Poland et al. (2012) for dense genome-wide coverage (Elshire et al., 2011). Markers with missing data greater than 50%, minor allele frequency less than 10% and pairwise marker correlation (r^2) greater than 0.95 (for redundancy) were filtered, that resulted in 5,102 markers for the 45th IBWSN, 8,066 markers for the 46th IBWSN and 8,857 markers for the combined nurseries. We also filtered for lines with greater than 50% missing data and obtained 267 lines, 305 lines and 566 lines for the 45th IBWSN, 46th IBWSN and the combined nurseries, respectively. The lines in the 45th IBWSN were genotyped using both GBS and DArTseq platforms. After using the same filtering criteria as above, we obtained 5,209 DArTseq markers for the 267 lines that in combination with 5,102 GBS markers resulted in 10,311 markers for the combined marker set. The expectation-maximization algorithm was implemented in the 'R' package rrBLUP (Endelman, 2011) to impute missing data.

Relationship matrix, linkage disequilibrium and heritability estimation

The genomic relationship matrix (G-matrix) was obtained according to VanRaden (2008) and implemented in the 'R' package rrBLUP (Endelman, 2011). It was centered and standardized for all the analyses. The linkage disequilibrium (LD) across the wheat chromosomes was

calculated using a subset of markers that were mapped. This comprised 3,531 GBS markers and 4,793 DArTseq markers for the 45th IBWSN. We used the r^2 measure of linkage disequilibrium which is the square of the correlation coefficient between two loci. It measures the proportion of the variance of a response variable that is explained by a predictor variable (Hill and Robertson, 1968). The value of r is given by,

$$r = \frac{D}{(p_1 p_2 q_1 q_2)^{1/2}}$$

where 'D' is the disequilibrium and p_1, q_1, p_2 and q_2 are the allele frequencies at the two loci. Heritability for the different traits was calculated on a line mean basis. The average information-restricted maximum likelihood algorithm (Gilmour et al., 1995) implemented in the 'heritability' package in 'R' (Kruijer et al., 2015) was used to obtain estimates of the genetic and residual variances.

Prediction models

Least-squares (LS)

A step-wise least squares (LS) approach was used that involved selection of markers and estimation of effects for the selected markers. First, a genome wide association analysis was conducted in the training set to identify the markers significantly associated with the trait. The markers were ranked according to their p-values for variable selection. For each iteration i through j, a marker was added to the model starting from the most significant marker,

$$\mathbf{y} = 1_n\mu + \mathbf{X}_i\beta_i \ldots\ldots \mathbf{X}_j\beta_j + \boldsymbol{\varepsilon}$$

where **y** is the phenotype, μ is the mean, β_i denotes the effect of the ith marker and $\mathbf{X}_i$ denotes the

i^{th} marker's genotype matrix. The 5-fold cross validation accuracy was calculated within the training set after each iteration and the model with j-1 markers was selected when the $Accuracy_{j-1} > Accuracy_j$. But two closely linked markers having very similar p-values will follow each other in the iteration and adding the second marker will not improve the accuracy. Since this would stop the iteration and lead to the exclusion of other linked markers with lower p-values, we removed markers that had pairwise marker correlation (r^2) greater than 0.80 for this model. Model selection was followed by estimation of marker effects from the selected model that were then used to predict the BVs of the individuals.

Genomic-best linear unbiased prediction

The genomic-best linear unbiased prediction **(**GBLUP) is a whole-genome regression approach that uses the G-matrix calculated from markers and has been successfully applied to predict complex traits (Yang et al., 2010; de los Campos et al., 2013b; Habier et al., 2013). It is equivalent to the ridge-regression best linear unbiased prediction (RR-BLUP) (Hoerl and Kennard, 1970; Whittaker et al., 2000; Piepho, 2009) when the similarity between the lines in genomic space is proportional to their genetic covariance (Habier et al., 2007; Goddard, 2009; Piepho, 2009; Endelman, 2011). The mixed model used in GBLUP to calculate BVs of individuals is,

$$\mathbf{y}=\mathbf{1}_n\boldsymbol{\mu}+\mathbf{Zu}+\boldsymbol{\varepsilon}$$

where $\mathbf{y}$ is the vector of the response phenotypic trait, $\boldsymbol{\mu}$ is the mean vector, $\mathbf{u}$ is the vector of genotype effects that are assumed to be multivariate normal random effects ($\mathbf{u} \sim N(0, G\sigma^2_u)$), $\mathbf{Z}$ is the design matrix for the random effects and $\boldsymbol{\varepsilon}$ is the vector of independent residuals assumed to

have a multivariate normal distribution ($\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^2_e)$). It was implemented using the 'R' package, rrBLUP (Endelman, 2011).

Bayesian models

The Bayesian models assume a prior marker effects distribution and are of the form:

$$\mathbf{y} = \mathbf{1}_n\boldsymbol{\mu} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where **X** is the incidence matrix for the markers and **β** is the vector of k marker effects. All the Bayesian models were implemented in the 'R' package BGLR (Pérez and de Los Campos, 2014). The default prior parameters were used with 50,000 iterations and the first 5,000 iterations were discarded as burn-in.

Bayesian ridge-regression

The BRR is the Bayesian counterpart of the RR-BLUP where the estimates of all the marker effects are shrunken towards zero. The shrinkage is independent of the effect size but dependent on the frequency and the sample size (Gianola, 2013). BRR is equivalent to the RR-BLUP, but instead of choosing the ridge parameter, a Gaussian prior that is independent and identically distributed (IID) with variance common to all the marker effects is used i.e. $p\left(\boldsymbol{\beta_R}\middle|\sigma^2_{\beta_R}\right) = \prod_{j=1}^{n} N\left(\boldsymbol{\beta_{R_j}}|0, \sigma^2_{\beta_R}\right)$ where $\boldsymbol{\beta_R}$ is the vector of regression coefficients and $\sigma^2_{\beta_R}$ is the *a priori* variance of marker effects (Perez et al., 2010). Then, the variance parameter ($\sigma^2_{\beta_R}$) was assigned a scaled-inverse Chi-squared density, $p(\sigma^2_{\beta_R}) = \chi^{-2}(\sigma^2_{\beta_R}|df_{\beta_R}, S_{\beta_R})$, where df_{β_R} and S_{β_R} are the prior degrees of freedom and scale, respectively. The default degrees of

freedom parameter was used ($df_\beta = 5$) and $S_{\beta_R}$was solved by BGLR to match the model R-squared.

Bayes A

As markers may contribute differentially to the genetic variance and the assumption of a common variance to all the markers may not be realistic. So, Meuwissen et al. (2001) proposed the BA model which induces marker-effect specific shrinkage. The scaled-t prior density is assigned to the marker effects which shrinks more strongly the markers with effects closer to zero but does not penalize severely the markers with large effects (Xu, 2003; Gianola, 2013). This density is implemented in BGLR as a mixture of scaled-normal densities. The BA hierarchical model involves two stages: in the first stage of hierarchy, normal densities with mean zero and variance parameters that are marker-specific ($\sigma^2{}_{\beta_{jk}}$) are assigned to the marker effects. In the second stage, IID scaled-inverse Chi-squared densities with known df_β and S_β are assigned to the marker variances (Gianola, 2013). The default degrees of freedom ($df_\beta = 5$) was used and the scale parameter (S_β) was assigned a gamma density (G) with rate parameter (r) and shape parameter (s). The prior densities for this model is represented as $\mathrm{p}\left(\boldsymbol{\beta_j}, \sigma^2{}_{\beta_j}, S_\beta\right) = \left\{\prod_k N\ (\boldsymbol{\beta_{jk}}|0, \sigma^2{}_{\beta_{jk}})\ \chi^{-2}\ (\sigma^2{}_{\beta_{jk}}|df_\beta, S_\beta)\right\} \mathrm{G}\ (S_\beta|r, s)$ (Pérez and de Los Campos, 2014).

Bayes B

Bayes B, also proposed by Meuwissen et al. (2001) uses a mixture distribution prior where marker effects are assumed to be zero with probability, π and marker effects are assumed to be

drawn from a scaled-t distribution with probability, 1-π. In BA, π = 0, but BB assumes that many markers have no effect at all and hence π > 0 (Habier et al., 2011). Heffner et al. (2009) referred to this as a more realistic prior because certain regions of the genome are expected to have no quantitative trait loci (QTL) and thereby zero effect. While zeroing out marker effects might not be ideal for infinitesimal traits, Meuwissen et al. (2001) argued that genetic variances are distributed across loci such that only a few have genetic variance and eliminating the marker effects close to zero reduces the noise. BGLR treats the parameter π (proportion of non-null effects) as unknown and assigns a Beta (B) prior parametrized such that the expected value by $E(\pi) = \pi_0$ and p_0 is the number of prior counts. The prior densities for BB is represented as $p(\boldsymbol{\beta_j}, \sigma_\beta^2, \pi) = \left\{\prod_k \left[\pi N\left(\boldsymbol{\beta_{jk}}|0, \sigma_\beta^2\right) + (1-\pi)1(\boldsymbol{\beta_{jk}} = 0)\right] \chi^{-2}\left(\sigma_{\beta_{jk}}^2 | df_\beta, S_\beta\right)\right\} B(\pi|p_0, \pi_0)$ x $G(S_\beta|r, s)$ (Pérez and de Los Campos, 2014).

Bayes Cπ

The Bayes Cπ (BC) model that is an extension of the Bayes C model (Kizilkaya et al., 2010) is similar to BB, except that it uses a Gaussian distribution instead of the t-density distribution used by BB (Habier et al., 2011; Lorenz et al., 2011). BC was developed to address the drawbacks of the BA and BB models. It treats the probability of markers with a zero effect (π) as unknown and estimates it instead of assuming a fixed π, as this could affect the shrinkage of marker effects (Habier et al., 2011). Hence, Bayes C is thought to be more flexible in modeling traits that are oligogenic to polygenic (Lorenz et al., 2011). The BC model implemented in BGLR

is similar to BB, except that the variance parameter ($\sigma^2{}_\beta$) is estimated from the data, p ($\sigma^2{}_\beta$) = χ^{-2} ($\sigma^2{}_\beta | df_\beta, S_\beta$).

Bayesian least absolute shrinkage and selection operator

The classical least absolute shrinkage and selection operator **(**LASSO) (Tibshirani, 1996) and its Bayesian counterpart (Park and Casella, 2008) combine the features of both shrinkage and subset selection. BL uses the double exponential distribution prior that has thick tails and places a higher density at zero (Perez et al., 2010). This is implemented in BGLR as a mixture of scaled normal densities. The marker effects with large absolute values are shrunken less (de los Campos et al., 2009). Independent normal densities with mean zero and marker-specific variance parameter (τ_{jk}^2x σ_ε^2) are assigned to the marker effects. A scaled-inverse Chi-square density is assigned to the residual variance. The marker specific scale parameters, τ_{jk}^2 are assigned IID exponential densities with rate parameter ($\lambda^2/2$) that was set to the default type gamma in BGLR (Pérez and de Los Campos, 2014). The prior densities for BL is represented as $p\left(\boldsymbol{\beta_j}, \tau_j^2, \lambda^2 | \sigma_\varepsilon^2\right) = \left\{\prod_k N\left(\boldsymbol{\beta_{jk}} | 0, \tau_{jk}^2 \text{ x } \sigma_\varepsilon^2\right) Exp\left\{\tau_{jk}^2 | \frac{\lambda^2}{2}\right\}\right\}$ x $G\ (\lambda^2 |\ r, s)$.

Reproducing kernel Hilbert spaces

The RKHS semi-parametric approach for genomic prediction (Gianola 2006; Gianola and van Kaam 2008) and is expected to capture some non-additive effects as it does not assume linearity. The RKHS model using a Gaussian kernel is of the form:

$$\boldsymbol{y}_i = \boldsymbol{w}_i{}'\boldsymbol{\beta} + \boldsymbol{z}_i{}'\boldsymbol{u} + \sum_{j=1}^{n} exp\left[-\frac{(x_i - x_j)'(x_i - x_j)}{h}\right]\alpha_j + \boldsymbol{\varepsilon_i}$$

where x_i and x_j are the observed marker genotypes of individuals, $\mathbf{w_i}$ and $\mathbf{z_i}$ are the incidence vectors, $\boldsymbol{\beta}$ is the vector of location effects, $\mathbf{u}$ is the vector of additive genetic effects, α_j is the regression coefficient and $\boldsymbol{\varepsilon_i}$ is the error term [$\boldsymbol{\varepsilon_i} \sim N(0, I\sigma^2_e)$] (Gianola et al., 2006). The additive genetic effects $u \sim N(0, K\sigma^2_g)$, where K is the reproducing Gaussian kernel, $K(x_i,x_j) = exp\left[-\frac{(x_i-x_j)'(x_i-x_j)}{h}\right]$ and σ^2_g is the additive genetic variance. We used the BGLR package (Pérez and de Los Campos, 2014) to implement three RKHS models: RKHS markers (RKHS-M) with the G-matrix calculated from markers; RKHS pedigree (RKHS-P) with the pedigree relationship matrix and RKHS markers and pedigree (RKHS-MP) with two kernels comprising the marker and pedigree relationship matrices. These models were fitted with three arbitrarily chosen bandwidth parameters and the three accuracies were averaged.

Model comparisons and genotyping by sequencing marker platform comparisons

The Pearson's correlation between the observed and the cross-validated BVs (prediction accuracy) was used to assess the predictive ability of the models. The 10-folds cross-validation was used and the training set comprised 240 lines in the 45th IBWSN, 275 lines in the 46th IBWSN and 509 lines in the combined nurseries. To compare the BVs predicted by the different models, the Spearman's rank correlations between the BVs for all the prediction models and traits were calculated. We also performed a hierarchical clustering to assess the similarity between the models. The cross-validated BVs for all the six datasets were standardized to zero mean and unit variance. A Euclidean distance matrix between the prediction models was then obtained using the standardized BVs and averaged for all the datasets. This distance matrix was then used to perform a hierarchical clustering of the models based on the Ward's criterion and a dendrogram was

constructed. We compared the prediction accuracies estimated from GBS markers, DArTseq markers and a combined set including markers obtained from both platforms.

RESULTS

Phenotypic data analysis

The phenotypic distributions for all the diseases in the combined dataset are shown in Figure 2.1. The mean correlation for STB across the years was moderate (0.47). The mean correlation between seedling resistance to SNB and TS was also moderate (0.33). Tan spot seedling resistance and APR had a low (0.16) correlation, indicating that the genetic bases of seedling resistance and APR were different.

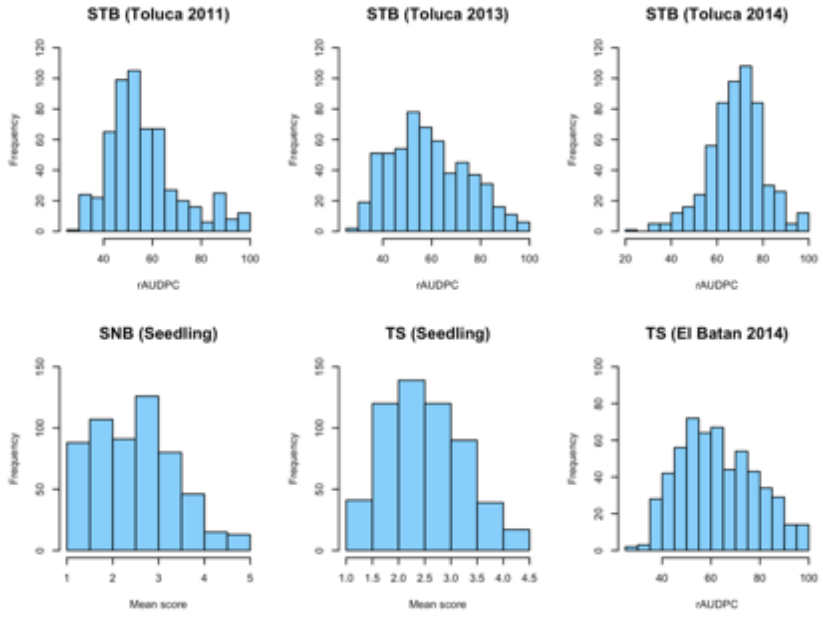

Figure 2.1: Phenotypic distributions for Septoria tritici blotch (STB), Stagonospora nodorum blotch (SNB) and tan spot (TS) in the 45th and 46th international bread wheat screening nursery (IBWSN) entries.

Relationship matrix and heritability estimation

Heatmap of the relationship matrices for the 566 lines in the 45th and 46th IBWSN (Figure 2.2) indicated that the pedigree relationship matrix shows more relatedness among the lines than the genomic relationship matrix. The 566 lines comprised 366 crosses which included one family with eight full-sibs, one with seven full-sibs, three with six full-sibs, one with five full-sibs, 14 with four full-sibs, 27 with three full-sibs, 72 with two full-sibs and 247 crosses represented by one individual per cross. The broad-sense, line-mean heritability was the highest for TS seedling (0.66), followed by SNB seedling (0.53) and STB APR (0.47). The broad sense heritability for TS APR was moderate (0.57).

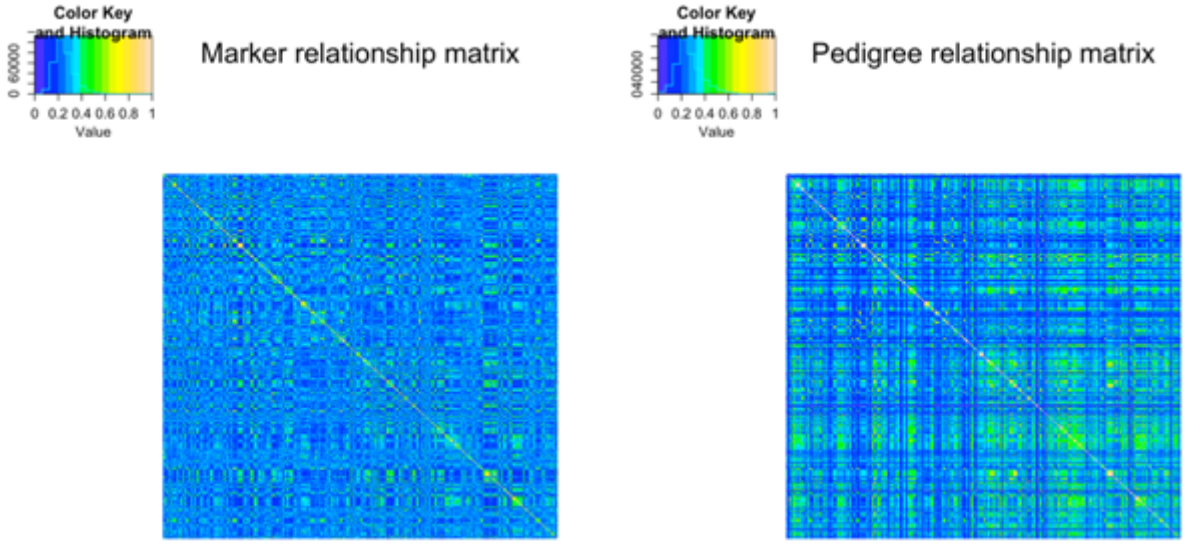

Figure 2.2: Heat map of the marker and pedigree based relationship matrices for the 566 lines in the 45th and 46th international bread wheat screening nursery (IBWSN) illustrating the familial relatedness (kinship) between the individuals.

Marker data and linkage disequilibrium

We observed that GBS generated higher missing data and had a higher number of markers with minor allele frequencies close to zero than DArTseq (Figure 2.3). Filtering for 50% missing data resulted in 13,913 GBS markers from 45,818 markers and 11,007 DArTseq markers from 11,211 markers.

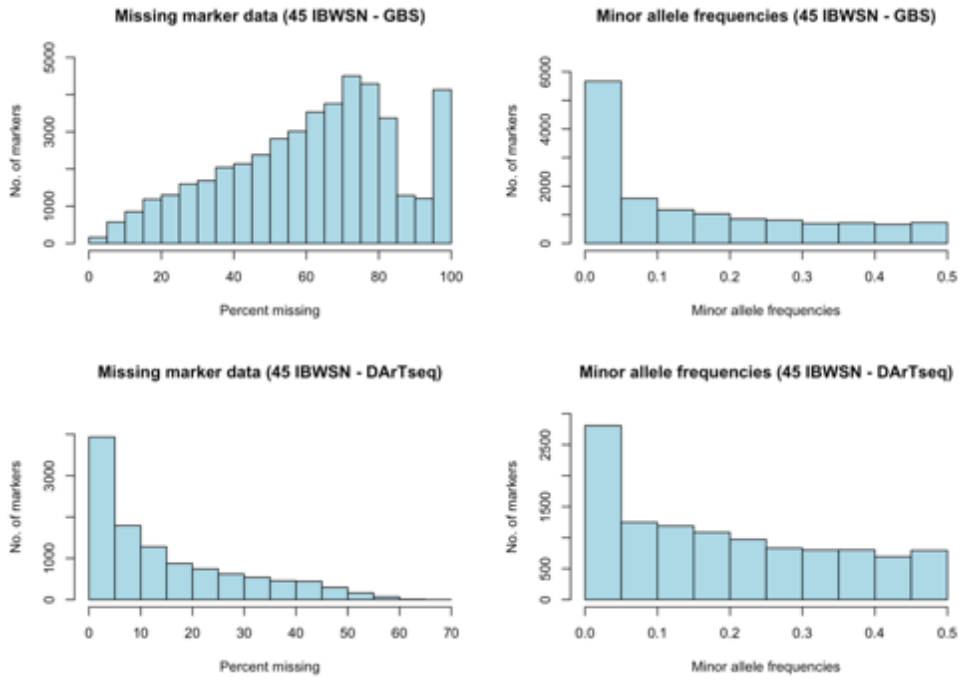

Figure 2.3: Missing marker data and minor allele frequencies of genotyping by sequencing (GBS) and diversity arrays technology-sequencing (DArTseq) markers.

We also compared the marker distributions across all the chromosomes (Figure 2.4) and found that chromosomes 2B and 4D had the highest and lowest proportion of GBS and DArTseq markers. The percentage of DArTseq markers on the A, B and D genomes were 46.25%, 46.35% and 7.4%. Similarly, the percentage of GBS markers on the A, B and D genomes were 39.9%, 50.1% and 10%. Overall, we did not observe significant differences in the percentage of marker coverage across the different chromosomes using these two whole-genome profiling approaches.

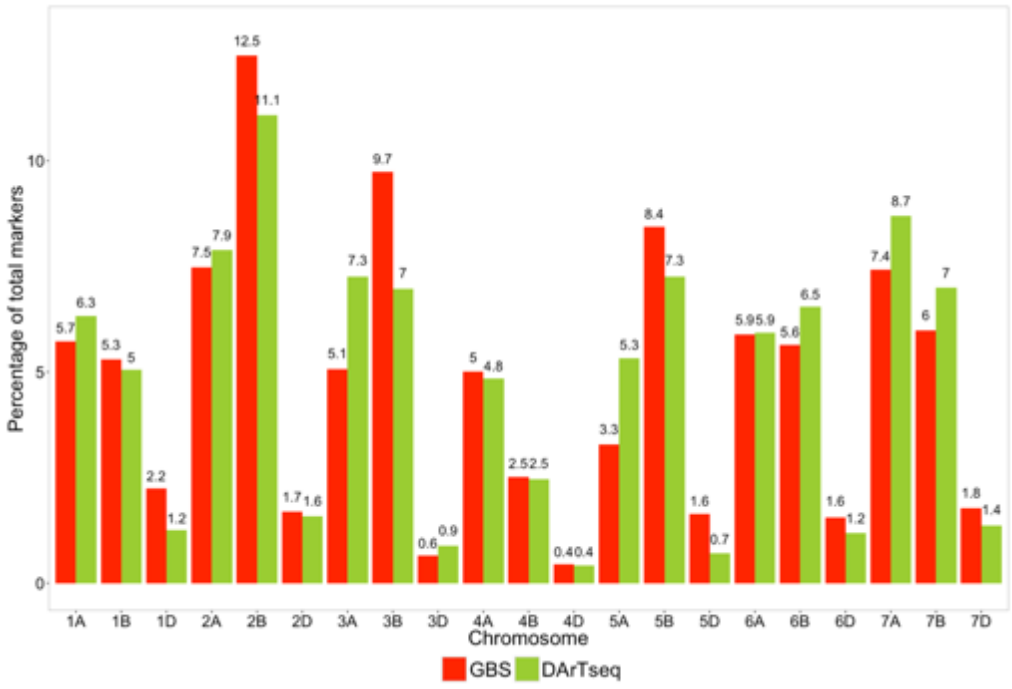

Figure 2.4: Distribution of genotyping by sequencing (GBS) and diversity arrays technology-sequencing (DArTseq) markers in the wheat chromosomes expressed as percentage of total markers.

The LD in the wheat chromosomes using the GBS and DArTseq markers (Figures 2.5 and 2.6) showed striking similarities across many chromosomes. Chromosomes 2A, 4A, 4B, 6B and 7B had large LD blocks which were observed using both the marker platforms. Although most of the D-genome chromosomes had similar LD patterns, it was hard to compare because of the limited number of markers. A few chromosomes (3A, 5A and 4B) had slightly different LD patterns using the two marker platforms, but this could be due to the differences in the number of the GBS and DArTseq markers used to calculate LD. On chromosome 1A, four small LD blocks were observed using GBS markers but not with the DArTseq markers, indicating clearly different LD patterns.

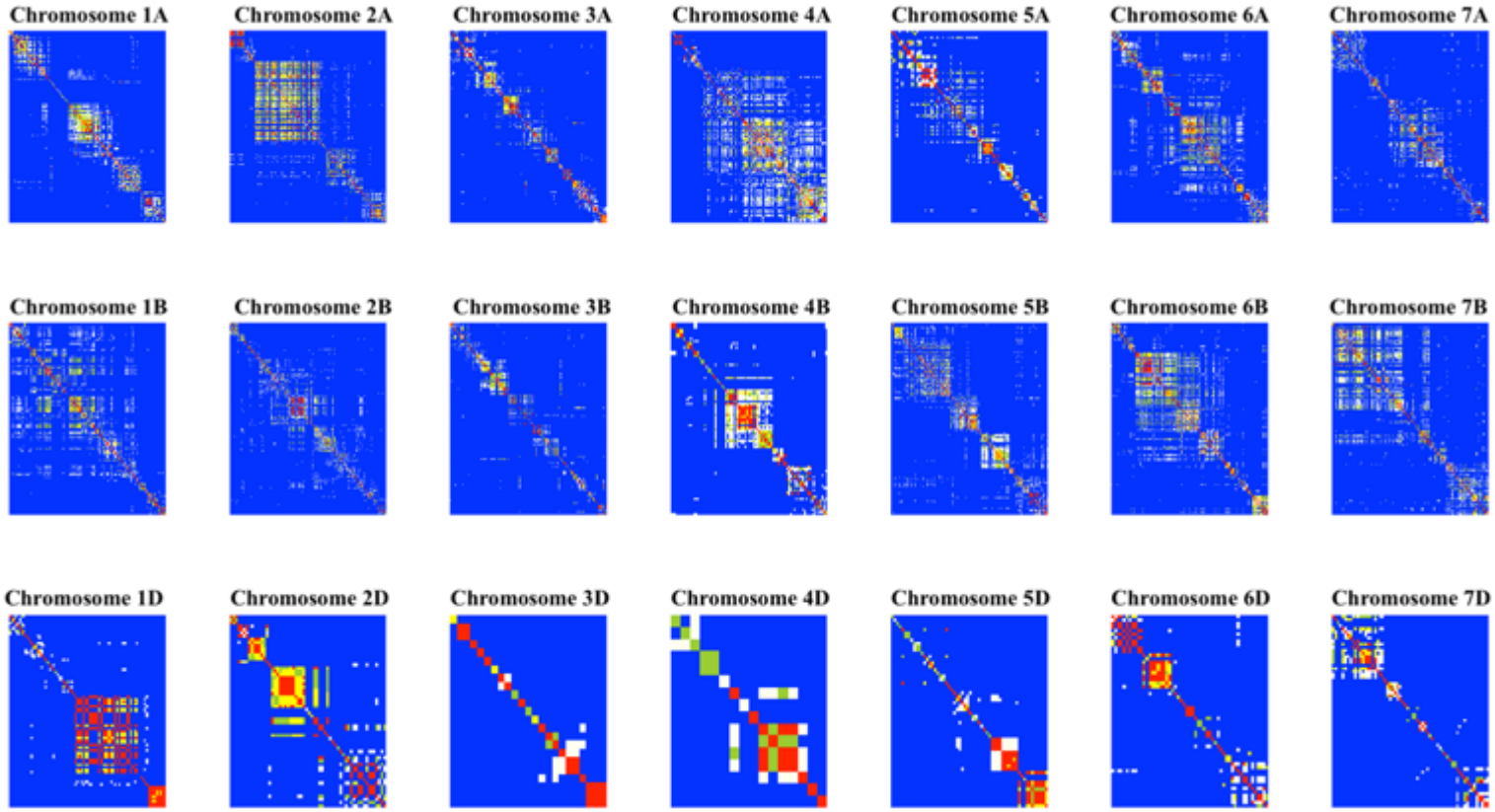

Figure 2.5: Linkage disequilibrium in the wheat chromosomes using the genotyping by sequencing (GBS) markers expressed as r^2 between marker loci. Blue represents r^2 of 0-0.2, white represents r^2 of 0.21-0.4, green represents r^2 of 0.41-0.6, yellow represents r^2 of 0.61-0.8 and red represents r^2 of 0.8-1.

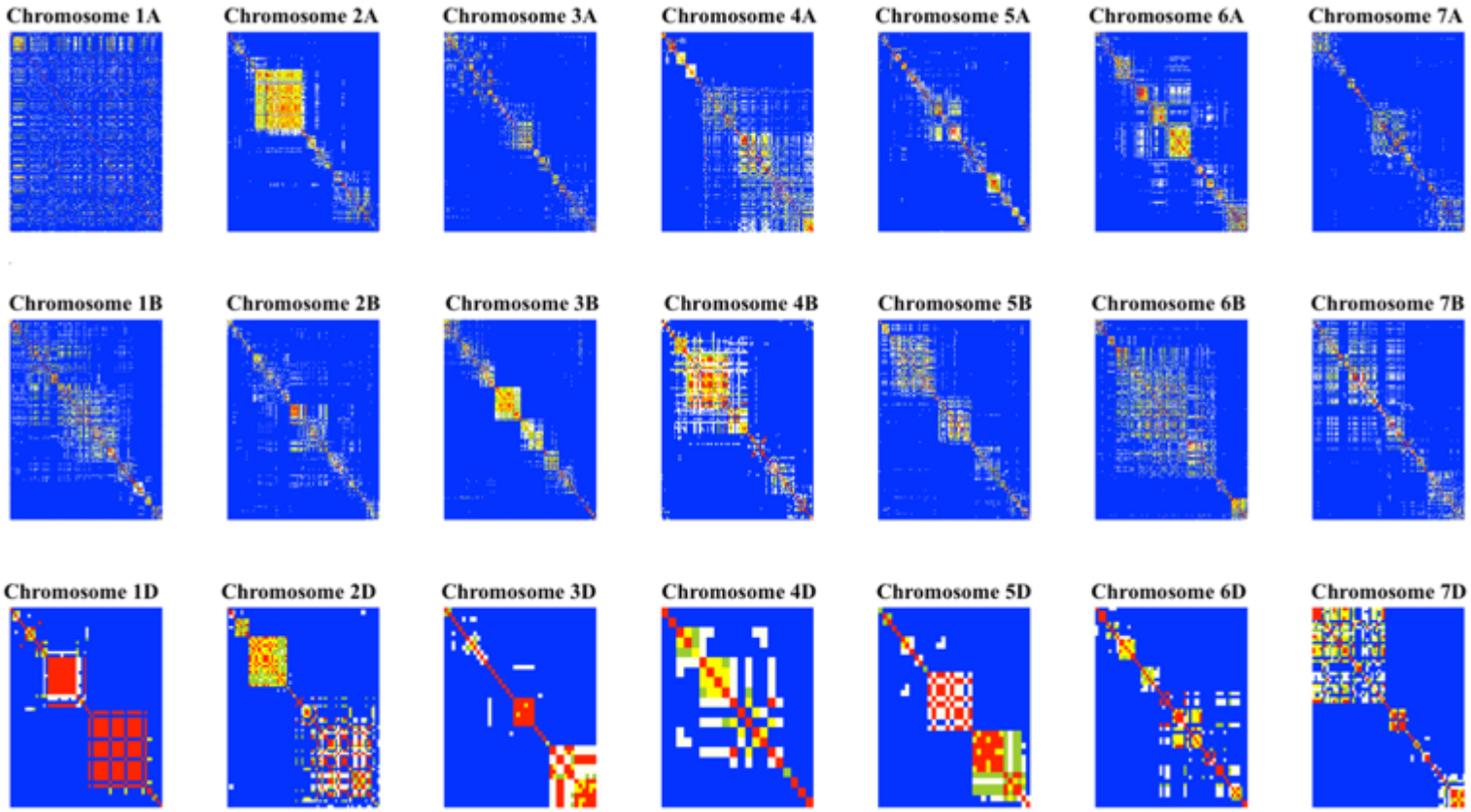

Figure 2.6: Linkage disequilibrium in the wheat chromosomes using the diversity arrays technology sequencing (DArTseq) markers expressed as r^2 between marker loci. Blue represents r^2 of 0-0.2, white represents r^2 of 0.21-0.4, green represents r^2 of 0.41-0.6, yellow represents r^2 of 0.61-0.8 and red represents r^2 of 0.8-1.

Prediction accuracies for Septoria tritici blotch adult plant resistance

For STB APR, in the 45th IBWSN, the models that gave the highest accuracies were: RKHS-MP in the 2011 dataset; both RKHS-P and the RKHS-MP in the 2013 dataset and RKHS-MP in the 2014 dataset (Table 2.1). The least accuracies were obtained using LS in all the datasets and the increase in accuracy using the genome-wide markers was 82%. The average number of markers used as fixed effects in the LS model were six (2011), five (2013) and three (2014). The most significant marker that occurred at a higher frequency in the folds explained 7.6%, 9.5% and 10% of the trait variation in the 2011, 2013 and 2014 datasets, respectively. The RKHS-P model performed similar to the genome-wide marker based models in all the datasets. In the 46th IBWSN, the models that gave the highest accuracies were: BB, BC and BL in the 2011 dataset; BB in the 2013 dataset and BA in the 2014 dataset. The genome-wide marker based models performed similar to LS in the 2014 dataset, but gave 95% increase in accuracy in the 2011 and 2013 datasets. The average number of markers used as fixed effects in the LS model were five (2013) and four (2011 and 2014). The most significant marker explained 9.2%, 8.1% and 13.1% of the trait variation in the 2011, 2013 and 2014 datasets, respectively. Genomic prediction models did slightly better than the RKHS-P in all the datasets (16% average increase in accuracy). In the combined dataset, the models that gave the highest accuracies were: RKHS-MP in both the 2011 and 2013 datasets and BB and BL in the 2014 dataset. The LS gave the lowest accuracies in all the datasets and the increase in accuracy using genome-wide marker based models was 71%. The average number of markers used as fixed effects in the LS model were three, five and six explaining 7.2%, 9.1% and 10% of the trait variation in the 2011, 2013 and 2014 datasets,

respectively. The RKHS-P model performed similar to the genomic prediction models in the 2011 and the 2013 datasets, but gave slightly lower accuracy in the 2014 dataset. Similar accuracies were obtained using GBLUP, BRR, BA, BB, BC, BL and RKHS-M models in the individual and combined nurseries across all the years. For the 2014 dataset, we also evaluated all the models without stripe rust as a covariate. This resulted in higher accuracies, due to a large effect stripe rust resistance locus that was highly significant in all the folds and explained >20% variation. Using stripe rust as a covariate accounted for this and the marker associated with stripe rust was no longer significant.

Prediction accuracies for Stagonospora nodorum blotch seedling resistance

For SNB seedling resistance, the RKHS-MP model performed the best both in the nurseries and in the combined dataset (Table 2.2). Genome-wide prediction models resulted in 36% increase in accuracy over LS in the 45th IBWSN and the combined dataset, but gave similar accuracies in the 46th IBWSN. The average number of markers used as fixed effects in the LS model was five in all the datasets. The most significant marker explained 15%, 17% and 16% of the variation in the 45th IBWSN, 46th IBWSN and the combined datasets, respectively. Similar accuracies were obtained with the pedigree model (RKHS-P) and all genome-wide marker based models (GBLUP, BRR, BA, BB, BC, BL and RKHS-M).

Prediction accuracies for tan spot seedling resistance

For TS seedling resistance, LS gave the lowest accuracies in all the datasets and the RKHS-MP model gave slightly higher accuracies (although it was not statistically different from the other

genome-wide marker based models) (Table 2.2). The increase in accuracy using genomic prediction models over LS was 48%. The average number of markers used as fixed effects in the LS model was four (45th IBWSN) and three (46th IBWSN and combined dataset). The most significant marker explained 23%, 18% and 17% of the variation in the 45th IBWSN, 46th IBWSN and the combined datasets, respectively. The RKHS-P model performed similar to the genome-wide marker based models in the 46th IBWSN and in the combined dataset. But in the 45th IBWSN, genomic prediction models performed slightly better than the pedigree (15.4% increase in accuracy). The accuracies obtained using GBLUP, BRR, BA, BB, BC, BL and RKHS-M models were similar.

Prediction accuracies for tan spot adult plant resistance

For TS APR, the RKHS-MP model gave the highest accuracy in the 45th IBWSN and the combined dataset. But GBLUP and BL models gave the highest accuracy in the 46th IBWSN. The LS gave the lowest accuracies and the increase in accuracy using genome-wide markers was 50%. The average number of markers used as fixed effects in the LS model were seven (45th IBWSN), one (46th IBWSN) and two (combined dataset). The most significant marker explained 11%, 15% and 16% of the variation the 45th, 46th IBWSN and the combined dataset, respectively. There were no significant differences in the accuracies obtained from GBLUP, BRR, BA, BB, BC, BL, RKHS-M and RKHS-P.

Table 2.1: Prediction accuracies for adult plant resistance (APR) to Septoria tritici blotch (STB) using different models in the 45th and 46th international bread wheat screening nurseries (IBWSN)

Model	STB (APR 2011)			STB (APR 2013)			STB (APR 2014)		
	45th	46th	45th and 46th	45th	46th	45th and 46th	45th	46th	45th and 46th
LS	0.24 ± 0.05	0.28 ± 0.06	0.22 ± 0.04	0.34 ± 0.05	0.19 ± 0.04	0.34 ± 0.05	0.20 ± 0.05	0.35 ± 0.05	0.28 ± 0.05
GBLUP	0.53 ± 0.04	0.43 ± 0.06	0.49 ± 0.04	0.46 ± 0.05	0.45 ± 0.04	0.49 ± 0.04	0.38 ± 0.07	0.39 ± 0.06	0.41 ± 0.04
BRR	0.53 ± 0.04	0.43 ± 0.06	0.48 ± 0.04	0.46 ± 0.05	0.45 ± 0.04	0.49 ± 0.04	0.37 ± 0.08	0.4 ± 0.06	0.40 ± 0.04
BA	0.53 ± 0.04	0.43 ± 0.07	0.49 ± 0.04	0.46 ± 0.06	0.47 ± 0.03	0.49 ± 0.04	0.38 ± 0.07	0.42 ± 0.06	0.41 ± 0.03
BB	0.53 ± 0.04	0.44 ± 0.07	0.49 ± 0.04	0.47 ± 0.06	0.48 ± 0.03	0.49 ± 0.04	0.37 ± 0.08	0.41 ± 0.06	0.42 ± 0.04
BC	0.53 ± 0.04	0.44 ± 0.06	0.49 ± 0.04	0.46 ± 0.05	0.46 ± 0.03	0.49 ± 0.04	0.39 ± 0.08	0.39 ± 0.06	0.41 ± 0.04
BL	0.51 ± 0.05	0.44 ± 0.06	0.48 ± 0.04	0.45 ± 0.05	0.46 ± 0.03	0.49 ± 0.04	0.38 ± 0.08	0.4 ± 0.06	0.42 ± 0.03
RKHS-M	0.53 ± 0.04	0.42 ± 0.06	0.49 ± 0.04	0.45 ± 0.05	0.46 ± 0.04	0.49 ± 0.04	0.39 ± 0.07	0.38 ± 0.06	0.41 ± 0.04
RKHS-P	0.54 ± 0.04	0.37 ± 0.05	0.48 ± 0.04	0.51 ± 0.04	0.41 ± 0.04	0.46 ± 0.04	0.37 ± 0.06	0.31 ± 0.04	0.34 ± 0.03
RKHS-MP	0.57 ± 0.03	0.42 ± 0.06	0.52 ± 0.03	0.51 ± 0.05	0.47 ± 0.04	0.50 ± 0.04	0.4 ± 0.07	0.39 ± 0.06	0.39 ± 0.03

APR, adult plant resistance; BA, Bayes A; BB, Bayes B; BC, Bayes Cπ; BL, Bayesian least absolute shrinkage and selection operator; BRR, Bayesian ridge regression; GBLUP, genomic best linear unbiased prediction; LS, least-squares; RKHS-M, reproducing kernel Hilbert spaces - markers; RKHS-P, reproducing kernel Hilbert spaces - pedigree; RKHS-MP, reproducing kernel Hilbert spaces - markers and pedigree; STB, Septoria tritici blotch.

Table 2.2: Prediction accuracies for Stagonospora nodorum blotch (SNB) and tan spot (TS) using different models in the 45th and 46th international bread wheat screening nurseries (IBWSN)

Model	SNB (seedling)			TS (seedling)			TS (APR)		
	45th	46th	45th and 46th	45th	46th	45th and 46th	45th	46th	45th and 46th
LS	0.43 ± 0.05	0.45 ± 0.04	0.43 ± 0.04	0.51 ± 0.05	0.43 ± 0.04	0.41 ± 0.03	0.28 ± 0.06	0.34 ± 0.04	0.35 ± 0.03
GBLUP	0.57 ± 0.04	0.49 ± 0.04	0.60 ± 0.02	0.76 ± 0.02	0.57 ± 0.03	0.66 ± 0.02	0.47 ± 0.05	0.42 ± 0.06	0.56 ± 0.03
BRR	0.58 ± 0.04	0.49 ± 0.04	0.60 ± 0.02	0.76 ± 0.02	0.56 ± 0.03	0.66 ± 0.02	0.48 ± 0.05	0.4 ± 0.06	0.56 ± 0.03
BA	0.57 ± 0.04	0.49 ± 0.04	0.60 ± 0.02	0.76 ± 0.03	0.56 ± 0.04	0.66 ± 0.02	0.48 ± 0.05	0.41 ± 0.06	0.56 ± 0.03
BB	0.57 ± 0.04	0.49 ± 0.04	0.59 ± 0.02	0.76 ± 0.02	0.56 ± 0.03	0.66 ± 0.02	0.47 ± 0.05	0.41 ± 0.06	0.56 ± 0.03
BC	0.58 ± 0.04	0.49 ± 0.04	0.59 ± 0.02	0.76 ± 0.02	0.56 ± 0.03	0.66 ± 0.02	0.48 ± 0.05	0.41 ± 0.06	0.56 ± 0.03
BL	0.57 ± 0.04	0.49 ± 0.04	0.60 ± 0.02	0.75 ± 0.02	0.56 ± 0.03	0.66 ± 0.02	0.48 ± 0.05	0.42 ± 0.06	0.56 ± 0.04
RKHS-M	0.58 ± 0.04	0.49 ± 0.04	0.59 ± 0.03	0.75 ± 0.02	0.56 ± 0.02	0.66 ± 0.02	0.47 ± 0.05	0.41 ± 0.06	0.56 ± 0.03
RKHS-P	0.55 ± 0.04	0.49 ± 0.03	0.60 ± 0.02	0.65 ± 0.03	0.55 ± 0.04	0.62 ± 0.03	0.46 ± 0.04	0.38 ± 0.04	0.52 ± 0.03
RKHS-MP	0.59 ± 0.03	0.52 ± 0.03	0.63 ± 0.02	0.77 ± 0.03	0.58 ± 0.04	0.68 ± 0.02	0.52 ± 0.05	0.40 ± 0.05	0.57 ± 0.03

APR, adult plant resistance; BA, Bayes A; BB, Bayes B; BC, Bayes Cπ; BL, Bayesian least absolute shrinkage and selection operator; BRR, Bayesian ridge regression; GBLUP, genomic best linear unbiased prediction; LS, least-squares; RKHS-M, reproducing kernel Hilbert spaces - markers; RKHS-P, reproducing kernel Hilbert spaces - pedigree; RKHS-MP, reproducing kernel Hilbert spaces - markers and pedigree; SNB, Stagonospora nodorum blotch; TS, tan spot.

Table 2.3: Spearman's rank correlations between estimated breeding values (BVs) for all the pairs of models

	LS	GBLUP	BRR	BA	BB	BC	BL	RKHS-M	RKHS-P	RKHS-MP
LS	1.00	0.57	0.57	0.55	0.56	0.56	0.57	0.57	0.46	0.54
GBLUP	0.57	1.00	1.00	0.99	1.00	1.00	0.99	1.00	0.68	0.91
BRR	0.57	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.68	0.91
BA	0.55	0.99	0.99	1.00	0.99	0.99	0.99	0.99	0.68	0.91
BB	0.56	1.00	0.99	0.99	1.00	0.99	0.99	0.99	0.68	0.91
BC	0.56	1.00	0.99	0.99	0.99	1.00	0.99	0.99	0.68	0.91
BL	0.57	0.99	0.99	0.99	0.99	0.99	1.00	0.99	0.68	0.91
RKHS-M	0.57	1.00	0.99	0.99	0.99	0.99	0.99	1.00	0.68	0.91
RKHS-P	0.46	0.68	0.68	0.68	0.68	0.68	0.68	0.68	1.00	0.87
RKHS-MP	0.54	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.87	1.00

BA, Bayes A; BB, Bayes B; BC, Bayes Cπ; BL, Bayesian least absolute shrinkage and selection operator; BRR, Bayesian ridge regression; GBLUP, genomic best linear unbiased prediction; LS, least-squares; RKHS-M, reproducing kernel Hilbert spaces - markers; RKHS-P, reproducing kernel Hilbert spaces - pedigree; RKHS-MP, reproducing kernel Hilbert spaces - markers and pedigree.

Comparisons between the prediction models

Overall, the RKHS-MP gave slightly higher accuracies and LS, the lowest across all the datasets. The accuracies obtained using GBLUP, BRR, BA, BB, BC, BL and RKHS-M models were similar. The RKHS-P model accuracies were not significantly different from the genome-wide marker based models. The average accuracies obtained using LS, genomic prediction models, RKHS-P and RKHS-MP were: 0.27, 0.45, 0.42 and 0.46, respectively for STB; 0.44, 0.55, 0.55 and 0.58, respectively for SNB; 0.45, 0.66, 0.61 and 0.68, respectively for TS (seedling) and 0.32, 0.48, 0.45 and 0.50, respectively for TS (APR). A cluster dendrogram with the hierarchical clustering of the prediction models based on cross-validated BVs (shown in Figure 2.7), makes it clear that the LS was different from the other models and branched out separately.

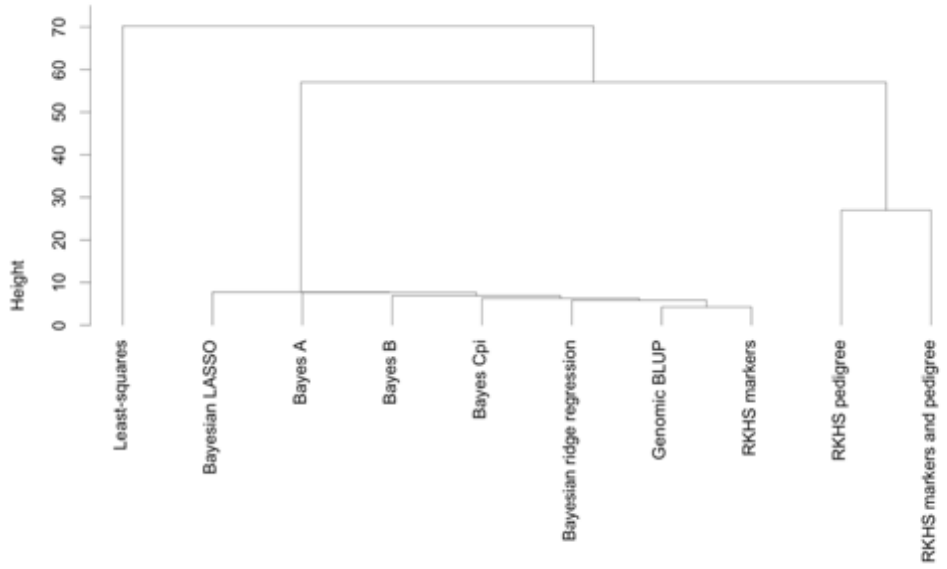

Figure 2.7: Cluster dendrogram showing the hierarchical clustering of the prediction models based on cross-validated estimated breeding values (BVs). LS, least-squares; RKHS, reproducing kernel Hilbert space; BLUP, best linear unbiased predictor; LASSO, least absolute shrinkage and selection operator.

RKHS-P and the RKHS-MP models clustered together. The BL, BA, BB, BC, BRR, GBLUP and RKHS-M clustered together. The Spearman's rank correlations between BVs for all the pairs of models (Table 3) shows that the BVs obtained from LS and RKHS-P had a moderate correlation with the BVs obtained from the genome-wide marker based models (0.57 and 0.68, respectively). The correlations among the BVs obtained from the other genome-wide marker based models were close to unity.

Comparisons between two whole-genome profiling approaches for genomic prediction

In the STB APR 2011 dataset, GBS markers, DArTseq markers and the combined marker set gave similar prediction accuracies using both LS and the genome-wide marker based models (Figure 2.8). In the STB APR 2013 dataset, GBS markers performed slightly better than the DArTseq markers with both LS (36% increase in accuracy) and genome-wide marker based models (21% increase in accuracy). The accuracies obtained using the combined marker set were similar to the accuracies obtained from the GBS markers. In the case of the STB APR 2014 dataset, GBS markers, DArTseq markers and the combined marker set gave similar accuracies with the LS. But with the genome-wide marker based models, GBS markers gave 36% increase in accuracy over DArTseq markers. The accuracies obtained using the combined marker set were not significantly different from the accuracies obtained from GBS markers. For SNB seedling resistance, GBS markers performed slightly better than the DArTseq markers with both LS (34% increase in accuracy) and genome-wide marker based models (36% increase in accuracy).

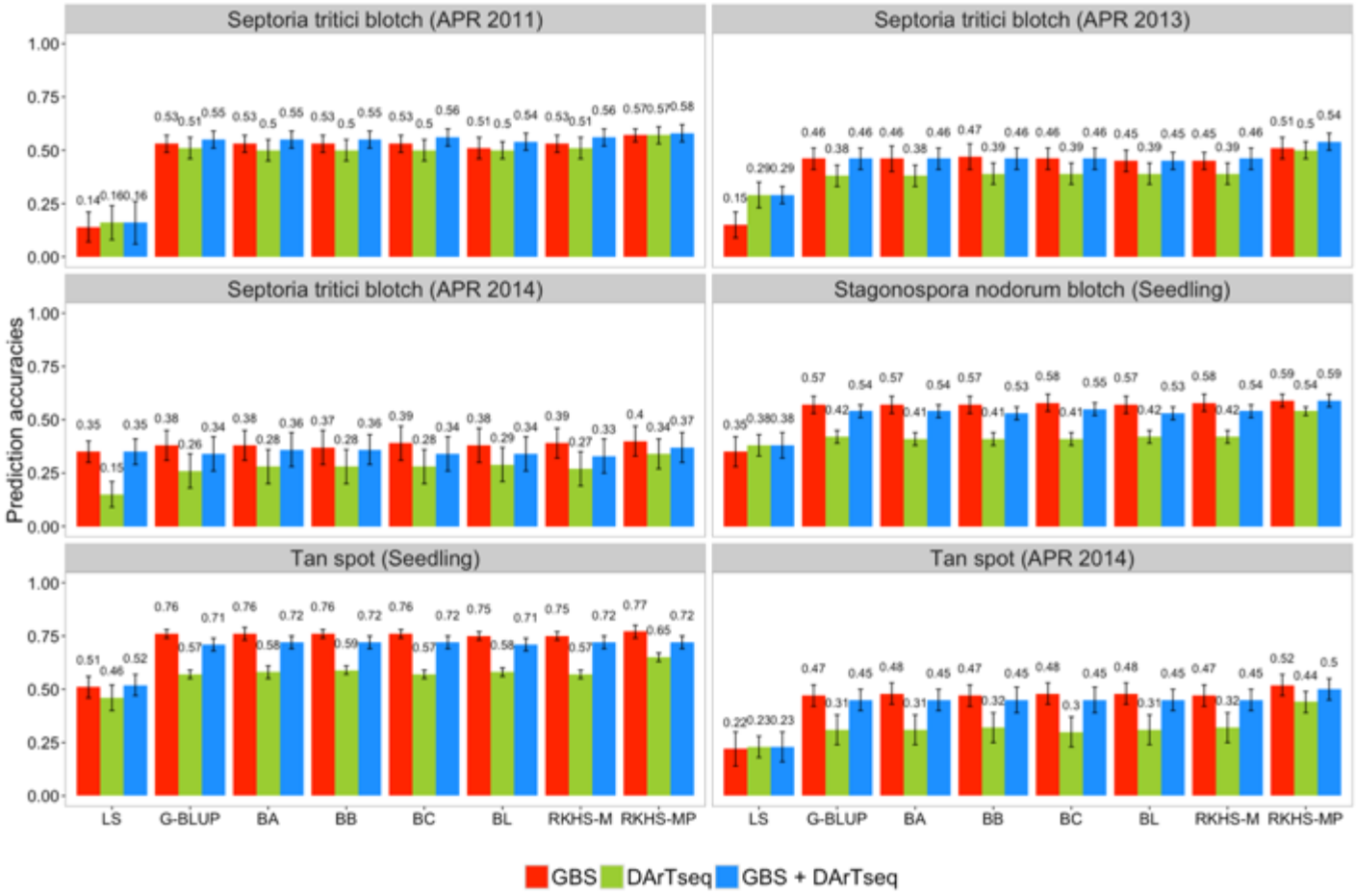

Figure 2.8: Prediction accuracies for Septoria tritici blotch, Stagonospora nodorum blotch and tan spot in the 45th international bread wheat screening nursery (IBWSN) using genotyping by sequencing (GBS), diversity arrays technology-sequencing (DArTseq) and both.

In the case of TS, DArTseq markers performed slightly better (18% increase in accuracy) for seedling resistance and GBS markers performed slightly better (33% increase in accuracy) for APR, using LS. But with the genome-wide marker based models, GBS markers outperformed the DArTseq-markers for both seedling resistance and APR (33.3% and 51.6% increase in accuracy, respectively). The combined marker set had slightly lower accuracies than the GBS markers with the genome-wide marker based models for both seeding resistance and APR, but the differences were not significant.

DISCUSSION

Among the diseases, the mean genomic prediction accuracies were the highest for seedling resistance to TS (0.66) and SNB (0.55), followed by APR to TS (0.48) and STB (0.45). The same trend was also observed with the LS approach and the highest prediction accuracies were obtained for seedling resistance to TS (0.45) and SNB (0.44), followed by APR to TS (0.32) and STB (0.27). These results indicate that genomic prediction models perform better than the LS approach, which is consistent with several previous studies (Meuwissen et al., 2001; Bernardo and Yu, 2007; Habier et al., 2007; Muir, 2007; Piyasatian et al., 2007; Lorenzana and Bernardo, 2009; Moser et al., 2009; Heffner et al., 2011a; b; Rutkoski et al., 2014). The average increase in accuracy using genomic prediction models compared to the LS approach was 48%. This is consistent with several previous reports: 41% (Meuwissen et al., 2001); 18 and 43% for a trait that has high and low heritability, respectively (Bernardo and Yu, 2007); 32% (Piyasatian et al., 2007) and 28% (Heffner et al., 2011a). Bernardo (2014) used simulations to show that the known QTLs can be fit as fixed effects only when they explain more than 10% of the genetic variance. In our study, the genetic variance

explained by the most significant marker ranged from 7% to 23%. For STB APR, the significant markers used as fixed effects explained less than 10% variation (except in one dataset), which indicates that resistance in these nurseries is quantitative and controlled by many genes with moderate to small effects. However, we also observed that for SNB and TS seedling resistance where the most significant marker explained greater than 10% variation, there were significant differences between the accuracies of LS and GBLUP indicating that in addition to the large effect loci, there were also minor loci controlling these traits. These results clearly demonstrate the advantage of using genome-wide markers for complex traits that are controlled by several QTL and support the infinitesimal model of Fisher (1918).

The GBLUP and BRR models resulted in accuracies similar to the other Bayesian models, despite the assumption that all the marker effects have equal variance. The use of different prior distributions for the marker effects in the Bayesian models did not affect the prediction accuracies. This is consistent with several previous studies that report similarities between these models for different traits: RR-BLUP and Bayesian regression (same as BA) for two traits in dairy bulls (Moser et al., 2009); BA, BB and BC using simulated and real data (Habier et al., 2011); RR-BLUP and BC for quantitative traits in elite North American oats (Asoro et al., 2011); RR, BA, BB and BC for several traits in wheat (Heffner et al., 2011a); RR-BLUP, BC and BL for Fusarium head blight resistance in barley (Lorenz et al., 2012); RR and BL for stem and stripe rust resistance in wheat (Ornella et al., 2012); GBLUP, BC and BL for stem rust APR resistance in wheat (Rutkoski et al., 2014). However, few studies have reported slightly higher accuracies with some Bayesian models. Some of the models that showed slight superiority over others are: BA and BB over BLUP in simulations (9% and 16% increase in accuracy, respectively) (Meuwissen et al., 2001); BB over GBLUP in simulations (Clark et al., 2011); BC and BA over RR-BLUP and BL

for Fusiform rust resistance in Loblolly pine (Resende et al., 2012); RR-BLUP and BB over BA and BC models for predicting hybrid wheat performance (Zhao et al., 2013). Few studies have also reported the superiority of GBLUP over models that use different prior distributions especially when the trait was controlled by a few QTL with large effects (Luan et al., 2009; Zhong et al., 2009; Daetwyler et al., 2010). However, for the traits that we analyzed, the equal variance assumption still holds good and the differential shrinkage of the Bayesian models which involves higher computational time might be unnecessary.

The RKHS-M model performed similar to GBLUP in our study. While some studies (Gianola et al., 2006; Crossa et al., 2010; Howard et al., 2014) have reported that the non-parametric models performed better than the parametric ones, Crossa et al. (2013) concluded that there was no clear superiority of either of the models. An interesting observation was that the RKHS-P did very well and markers only gave 5.6% improvement in overall accuracies. This is consistent with several studies that have reported slight superiority of marker based models over the pedigree (Crossa et al. 2010; Spindel et al. 2015). However, the genomic-based relationship is expected to predict the allele sharing (within family variation) or the Mendelian sampling with better accuracy (Villanueva et al., 2005; Daetwyler et al., 2007; Goddard and Hayes, 2007). Some of the benefits of using the G-matrix include: avoiding selection of closely related sibs (Daetwyler et al., 2007), providing better accuracies when unrelated individuals are involved (van der Werf, 2009) and correcting for pedigree errors (Munoz et al., 2014). The high accuracies obtained with the pedigree in our study can be due to the following: the excellent pedigree recording system at CIMMYT that goes back several generations, small family sizes that have a minimal Mendelian sampling component for markers and the inclusion of full-sibs in both the training and validation sets. But it should be noted that this resulted from the use of late generation lines and might not

work as well for unselected lines in early generations. We also observed that the RKHS-MP model performed better than just the pedigree (9.9% increase in accuracy) and markers (3.6% increase in accuracy) alone and gave the highest accuracies for most datasets. This is consistent with several previous studies (de los Campos et al., 2009; Crossa et al., 2010b, 2013; Burgueño et al., 2012). Thus, the pedigree in conjunction with molecular markers can enhance the accuracy of selections.

Our comparisons between the LD captured by the two whole-genome profiling approaches, GBS and DArTseq indicated that they were similar except for a few chromosomes. For predictions using the LS approach, the accuracies were similar for GBS and DArTseq in the STB 2014 dataset and DArTseq performed better than GBS for TS seedling resistance (17.6% increase in accuracy). But, GBS performed slightly better than DArTseq in all the other datasets resulting in 34% mean increase in accuracy. Similarly, GBS performed slightly better (28.4% increase in mean accuracy) than DArTseq for all the diseases using genomic prediction models. This is consistent with a previous study by Heslot et al. (2013) who obtained a higher accuracy using GBS markers compared to DArT markers. We attribute our results to the following: (i) Both the approaches use different restriction enzymes for complexity reduction. While GBS uses the combination of *PstI* and *MspI* (Poland et al., 2012), DArTseq uses two complexity reduction methods with *PstI/HpaII* and *PstI/HhaI* followed by selection of a subset of fragments (Sansaloni et al., 2011; Li et al., 2015). (ii) DArTseq is done at a higher sequencing depth and uses strict filtering criteria that generates markers with less missing data compared to GBS. But this could also lead to the loss of some rare informative markers. (iii) Although it was not possible to compare the differences in marker coverage across the genome using the two approaches (because the positions of all the markers were not available), inadequate marker coverage in regions associated with the trait could lead to higher accuracies with one approach over the other. An interesting observation was that the

combined marker set with both GBS and DArTseq markers did not improve the prediction accuracies. This might be due to the redundancy in information captured by both these marker platforms. However, our results could be specific to the population used and the traits considered. It might not be necessarily true for other populations and traits.

In conclusion, we have used a range of models including a variable selection method, shrinkage methods, kernel-based methods and two whole-genome profiling approaches to predict resistance to wheat leaf spotting diseases. Our results clearly indicate that using genomic prediction is advantageous to selecting based on a few markers in marker-assisted selection. While model choice and genotyping approach are key elements for implementing GS, the genetic architecture of the trait, heritability, marker density, LD between the QTL and the markers, training population size and the relatedness between the individuals in the training and validation populations also play an important role in making decisions (de Roos et al., 2009; Lorenzana and Bernardo, 2009; Luan et al., 2009; Daetwyler et al., 2010b; Clark et al., 2011; de los Campos et al., 2013a; Howard et al., 2014). We hope that implementing GS in breeding for complex leaf spotting disease resistance in wheat will result in higher accuracy and rapid gains from selection.

REFERENCES

Abeysekara, N.S., T.L. Friesen, B. Keller, and J.D. Faris. 2009. Identification and characterization of a novel host-toxin interaction in the wheat-*Stagonospora nodorum* pathosystem. Theor Appl Genet 120: 117–126.

Asoro, F.G., M.A. Newell, W.D. Beavis, M.P. Scott, and J.-L. Jannink. 2011. Accuracy and Training Population Design for Genomic Selection on Quantitative Traits in Elite North American Oats. Plant Genome 4: 132–144.

Basnet, B.R., R.P. Singh, S.A. Herrera-Foessel, A.M.H. Ibrahim, J. Huerta‐Espino, V. Calvo-Salazar, et al. 2013. Genetic Analysis of Adult Plant Resistance to Yellow Rust and Leaf Rust in Common Spring Wheat Quaiu 3. Plant Dis 97(6): 728–736.

Bernardo, R. 2014. Genomewide selection when major genes are known. Crop Sci 54: 68–75.

Bernardo, R., and J. Yu. 2007. Prospects for genomewide selection for quantitative traits in maize. Crop Sci 47: 1082–1090.

Bhathal, J.S., R. Loughman, and J. Speijers. 2003. Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and Septoria nodorum blotch. Eur J Plant Pathol 109: 435–443.

Bolton, M.D., J.A. Kolmer, and D.F. Garvin. 2008. Wheat leaf rust caused by Puccinia triticina. Mol Plant Pathol 9: 563–575.

Box, G.E.P., and D.R. Cox. 1964. An Analysis of Transformations. J R Stat Soc Ser B 26(2): 211–252.

Brading, P.A., E.C.P. Verstappen, G.H.J. Kema, and J.K.M. Brown. 2002. A Gene-for-Gene Relationship Between Wheat and *Mycosphaerella graminicola*, the Septoria Tritici Blotch Pathogen. Phytopathology 92(4): 439–445.

Brown, J.K.M., L. Chartrain, P. Lasserre-Zuber, and C. Saintenac. 2015. Genetics of resistance to *Zymoseptoria tritici* and applications to wheat breeding. Fungal Genet Biol 79: 33–41.

Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa. 2012. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Sci 52: 707–719.

Chapman, J.A., M. Mascher, A. Buluç, K. Barry, E. Georganas, A. Session, et al. 2015. A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome.

Genome Biol 16(1): 26.

Chen, X.M. 2005. Epidemiology and control of stripe rust [Puccinia striiformis f . sp . tritici] on wheat. Can J Plant Pathol 27: 314–337.

Clark, S.A., J.M. Hickey, and J.H. van der Werf. 2011. Different models of genetic variation and their effect on genomic evaluation. Genet Sel Evol 43: 18.

Crook, A.D., T.L. Friesen, Z.H. Liu, P.S. Ojiambo, and C. Cowger. 2012. Novel Necrotrophic Effectors from *Stagonospora nodorum* and Corresponding Host Sensitivities in Winter Wheat Germplasm in the Southeastern United States. Phytopathology 102: 498–505.

Crossa, J., Y. Beyene, S. Kassa, P. Perez, J.M. Hickey, C. Chen, et al. 2013. Genomic Prediction in Maize Breeding Populations with Genotyping-by-Sequencing. G3 Genes|Genomes|Genetics 3(11): 1903–1926.

Crossa, J., J. Burgueño, S. Dreisigacker, M. Vargas, S.A. Herrera-Foessel, M. Lillemo, et al. 2007. Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics 177(3): 1889–1913.

Crossa, J., G. de los Campos, P. Pérez, D. Gianola, J. Burgueño, J.L. Araus, et al. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186: 713–724.

Crossa, J., P. Pérez, J. Hickey, J. Burgueño, L. Ornella, J. Cerón-Rojas, et al. 2014. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity (Edinb) 112: 48–60.

Daetwyler, H.D., J.M. Hickey, J.M. Henshall, S. Dominik, B. Gredler, J.H.J. Van Der Werf, et al. 2010a. Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population. Anim Prod Sci 50: 1004–1010.

Daetwyler, H.D., R. Pong-Wong, B. Villanueva, and J.A. Woolliams. 2010b. The impact of

genetic architecture on genome-wide evaluation methods. Genetics 185: 1021–1031.

Daetwyler, H.D., B. Villanueva, P. Bijma, and J.A. Woolliams. 2007. Inbreeding in genome-wide selection. J Anim Breed Genet 124: 369–376.

Dekkers, J.C.M., and F. Hospital. 2002. The Use of Molecular Genetics in the Improvement of Agricultural Populations. Nat Rev Genet 3(1): 22–32.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, et al. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS One 6(5): e19379.

Endelman, J.B. 2011. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. Plant Genome 4: 250–255.

Eyal, Z., A.L. Scharen, J.M. Prescott, and M. van Ginkel. 1987. The Septoria Diseases of Wheat: Concepts and methods of disease management. CIMMYT., Mexico, D.F.

Faris, J.D., Z. Liu, and S.S. Xu. 2013. Genetics of tan spot resistance in wheat. Theor Appl Genet 126: 2197–2217.

Feng, J., H. Ma, and G.R. Hughes. 2004. Genetics of resistance to Stagonospora nodorum blotch of hexaploid wheat. Crop Sci 44: 2043–2048.

Fisher, R.A. 1918. The Correlation between Relatives on the Supposition of Mendelian Inheritance. Trans R Soc Edinburgh 52: 399–433.

Flor, H.H. 1956. The complementary genic systems in flax and flax rust. Adv Genet 8: 29–54.

Fones, H., and S. Gurr. 2015. The impact of Septoria tritici Blotch disease on wheat: An EU perspective. Fungal Genet Biol 79: 3–7.

Friesen, T.L., C. Chu, S.S. Xu, and J.D. Faris. 2012. SnTox5-*Snn5*: A novel *Stagonospora nodorum* effector-wheat gene interaction and its relationship with the SnToxA-*Tsn1* and

SnTox3-*Snn3-B1* interactions. Mol Plant Pathol 13: 1101–1109.

Friesen, T.L., S.W. Meinhardt, and J.D. Faris. 2007. The *Stagonospora nodorum*-wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. Plant J 51: 681–692.

Friesen, T.L., E.H. Stukenbrock, Z. Liu, S. Meinhardt, H. Ling, J.D. Faris, et al. 2006. Emergence of a new disease as a result of interspecific virulence gene transfer. Nat Genet 38(8): 953–956.

Friesen, T.L., Z. Zhang, P.S. Solomon, R.P. Oliver, and J.D. Faris. 2008. Characterization of the interaction of a novel *Stagonospora nodorum* host-selective toxin with a wheat susceptibility gene. Plant Physiol 146: 682–693.

Gao, Y., J.D. Faris, Z. Liu, Y.M. Kim, R.A. Syme, R.P. Oliver, et al. 2015. Identification and Characterization of the SnTox6-Snn6 Interaction in the Parastagonospora nodorum – Wheat Pathosystem. Mol Plant-Microbe Interact 28(5): 615–625.

Ghaffary, S.M.T., J.D. Faris, T.L. Friesen, R.G.F. Visser, T.A.J. van der Lee, O. Robert, et al. 2012. New broad-spectrum resistance to septoria tritici blotch derived from synthetic hexaploid wheat. Theor Appl Genet 124: 125–142.

Gianola, D. 2013. Priors in whole-genome regression: The Bayesian alphabet returns. Genetics 194(3): 573–596.

Gianola, D., R.L. Fernando, and A. Stella. 2006. Genomic-Assisted Prediction of Genetic Value With Semiparametric Procedures. Genetics 173: 1761–1776.

Gianola, D., and J.B.C.H.M. van Kaam. 2008. Reproducing Kernel Hilbert Spaces Regression Methods for Genomic Assisted Prediction of Quantitative Traits. Genetics 178(4): 2289–2303.

Gilchrist-Saavedra, L., G. Fuentes-Dávila, C. Martínez-Cano, R.M. López-Atilano, E. Duveiller, R.P. Singh, et al. 2006. Practical guide to the identification of selected diseases of wheat and barley. Second. CIMMYT, Mexico, D.F.

Gilmour, A.R., R. Thompson, and B.R. Cullis. 1995. Average Information REML: An Efficient Algorithm for Variance Parameter Estimation in Linear Mixed Models. Biometrics 51(4): 1440–1450.

van Ginkel, M., and S. Rajaram. 1993. Breeding for durable resistance to diseases in wheat an additional perspective. Durab Dis Resist: 259–272.

Goddard, M. 2009. Genomic selection: prediction of accuracy and maximisation of long term response. Genetica 136(2): 245–257.

Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. J Anim Breed Genet 124(6): 323–330.

Goodwin, S.B. 2007. Back to basics and beyond: Increasing the level of resistance to Septoria tritici blotch in wheat. Australas Plant Pathol 36: 532–538.

Habier, D., R.L. Fernando, and J.C.M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177(4): 2389–2397.

Habier, D., R.L. Fernando, and J.C.M. Dekkers. 2009. Genomic Selection Using Low-Density Marker Panels. Genetics 182(1): 343–353.

Habier, D., R.L. Fernando, and D.J. Garrick. 2013. Genomic BLUP decoded: A look into the black box of genomic prediction. Genetics 194(3): 597–607.

Habier, D., R.L. Fernando, K. Kizilkaya, and D.J. Garrick. 2011. Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics 12: 186.

Hayes, B., and M. Goddard. 2010. Genome-wide association and genomic selection in animal breeding. Genome 53(11): 876–83.

Hayes, B.J., P.M. Visscher, and M.E. Goddard. 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genet Res (Camb) 91(1): 47–60.

Heffner, E.L., J.L. Jannink, H. Iwata, E. Souza, and M.E. Sorrells. 2011a. Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci 51: 2597–2606.

Heffner, E.L., J. Jannink, and M.E. Sorrells. 2011b. Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. Plant Genome 4(1): 65–75.

Heffner, E.L., M.E. Sorrells, and J.-L. Jannink. 2009. Genomic Selection for Crop Improvement. Crop Sci 49: 1–12.

Heslot, N., D. Akdemir, M.E. Sorrells, and J.L. Jannink. 2013a. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor Appl Genet: 1–18.

Heslot, N., J.L. Jannink, and M.E. Sorrells. 2013b. Using genomic prediction to characterize environments and optimize prediction accuracy in applied breeding data. Crop Sci 53(June): 921–933.

Heslot, N., J.-L. Jannink, and M.E. Sorrells. 2015. Perspectives for Genomic Selection Applications and Research in Plants. Crop Sci 55(february): 1–12.

Heslot, N., J. Rutkoski, J. Poland, J.L. Jannink, and M.E. Sorrells. 2013c. Impact of Marker Ascertainment Bias on Genomic Selection Accuracy and Estimates of Genetic Diversity. PLoS One 8(9): e74612.

Heslot, N., H.-P. Yang, M.E. Sorrells, and J.-L. Jannink. 2012. Genomic Selection in Plant Breeding: A Comparison of Models. Crop Sci 52: 146–160.

Hill, W.G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor Appl Genet 38(6): 226–231.

Hoerl, A.E., and R.W. Kennard. 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12(1): 55–67.

Howard, R., A.L. Carriquiry, and W.D. Beavis. 2014. Parametric and Nonparametric Statistical Methods for Genomic Selection of Traits with Additive and Epistatic Genetic Architectures. G3 Genes|Genomes|Genetics 4(6): 1027–1046.

Huerta-Espino, J., R.P. Singh, S. Germán, B.D. McCallum, R.F. Park, W.Q. Chen, et al. 2011. Global status of wheat leaf rust caused by *Puccinia triticina*. Euphytica 179(1): 143–160.

Jiang, Y., and J.C. Reif. 2015. Modelling Epistasis in Genomic Selection. Genetics: genetics.115.177907-.

Johnson, R. 1984. A critical analysis of durable resistance. Annu Rev Phytopathol 22: 309–330.

Kizilkaya, K., R.L. Fernando, and D.J. Garrick. 2010. Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes. J Anim Sci 88: 544–551.

Kolmer, J.A., J.A. Anderson, and J.M. Flor. 2010. Chromosome location, linkage with simple sequence repeat markers, and leaf rust resistance conditioned by gene *Lr63* in wheat. Crop Sci 50(6): 2392–2395.

Kruijer, W., M.P. Boer, M. Malosetti, P.J. Flood, B. Engel, R. Kooke, et al. 2015. Marker-Based Estimation of Heritability in Immortal Populations. Genetics 199(2): 379–398.

Lamari, L., and C.C. Bernier. 1989a. Evaluation of wheat lines and cultivars to tan spot [*Pyrenophora tritici-repentis*] based on lesion type. Can J Plant Pathol 11: 49–56.

Lamari, L., and C.C. Bernier. 1989b. Virulence of isolates of *Pyrenophora tritici-repentis* on 11 wheat cultivars and cytology of the differential host reactions. Can J Plant Pathol 11(3): 284–290.

Lan, C.X., R.P. Singh, J. Huerta-Espino, V. Calvo-Salazar, and S.A. Herrera-Foessel. 2014. Genetic Analysis of Resistance to Leaf Rust and Stripe Rust in Wheat Cultivar Francolin#1. Plant Dis 98(9): 1227–1234.

Leonard, K.J., and L.J. Szabo. 2005. Stem rust of small grains and grasses caused by Puccinia graminis. Mol Plant Pathol 6(2): 99–111.

Li, H., P. Vikram, R.P. Singh, A. Kilian, J. Carling, J. Song, et al. 2015. A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. BMC Genomics 16: 216.

Liu, Z.H., J.D. Faris, S.W. Meinhardt, S. Ali, J.B. Rasmussen, and T.L. Friesen. 2004. Genetic and Physical Mapping of a Gene Conditioning Sensitivity in Wheat to a Partially Purified Host-Selective Toxin Produced by *Stagonospora nodorum*. Phytopathology 94(5): 1056–1060.

Liu, Z., T.L. Friesen, H. Ling, S.W. Meinhardt, R.P. Oliver, J.B. Rasmussen, et al. 2006. The *Tsn1*-ToxA interaction in the wheat-*Stagonospora nodorum* pathosystem parallels that of the wheat-tan spot system. Genome 49: 1265–1273.

Liu, Z., Z. Zhang, J.D. Faris, R.P. Oliver, R. Syme, M.C. McDonald, et al. 2012. The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring *Snn1*. PLoS Pathog 8(1): e1002467.

Lopez-Cruz, M., J. Crossa, D. Bonnett, S. Dreisigacker, J. Poland, J.-L. Jannink, et al. 2015. Increased Prediction Accuracy in Wheat Breeding Trials Using a Marker x Environment Interaction Genomic Selection Model. G3 Genes|Genomes|Genetics 5(4): 569–82.

Lorenz, A.J., S. Chao, F.G. Asoro, E.L. Heffner, T. Hayashi, H. Iwata, et al. 2011. Genomic Selection in Plant Breeding: Knowledge and prospects. Adv Agron 110: 77–123.

Lorenzana, R.E., and R. Bernardo. 2009. Accuracy of genotypic value predictions for marker-

based selection in biparental plant populations. Theor Appl Genet 120(1): 151–161.

Lorenz, A.J., K.P. Smith, and J.L. Jannink. 2012. Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. Crop Sci 52: 1609–1621.

De los Campos, G., D. Gianola, G.J.M. Rosa, K. a Weigel, and J. Crossa. 2010. Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. Genet Res (Camb) 92(4): 295–308.

de los Campos, G., J.M. Hickey, R. Pong-Wong, H.D. Daetwyler, and M.P.L. Calus. 2013a. Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. Genetics 193: 327–345.

de los Campos, G., H. Naya, D. Gianola, J. Crossa, A. Legarra, E. Manfredi, et al. 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics 182: 375–385.

de los Campos, G., A.I. Vazquez, R. Fernando, Y.C. Klimentidis, and D. Sorensen. 2013b. Prediction of complex human traits using the genomic best linear unbiased predictor. PLoS Genet 9(7): e1003608.

Luan, T., J.A. Woolliams, S. Lien, M. Kent, M. Svendsen, and T.H.E. Meuwissen. 2009. The accuracy of genomic selection in Norwegian red cattle assessed by cross-validation. Genetics 183: 1119–1126.

Marasas, C.N., M. Smale, and R.P. Singh. 2004. The Economic Impact in Developing Countries of Leaf Rust Resistance Breeding in CIMMYT-Related spring bread wheat. CIMMYT Mex DF.

McIntosh, R.A., J. Dubcovsky, W. Rogers, C. Morris, R. Appels, and X.C. Xia. 2016. Catalogue of gene symbols for wheat: 2015-2016 supplement.

McIntosh, R., C. Wellings, and R. Park. 1995. Wheat rusts: an atlas of resistance genes. CSIRO Publishing.

McNeal, F.H., C.F. Konzak, E.P. Smith, W.S. Tate, and T.S. Russell. 1971. A uniform system for recording and processing cereal research data. Agricultural Research Service, United States Department of Agriculture, [Beltsville Md.].

Meien-Vogeler, F., G. Bartels, and H. Fehrmann. 1994. Distribution of *S. tritici* on wheat in Germany and its regional importance. p. 299–300. *In* Proceedings of the 4th International Septoria of Cereals Workshop. IHAR, Radzikow, Poland.

Meuwissen, T. 2007. Genomic selection : marker assisted selection on a genome wide scale. J Anim Breed Genet 124: 321–322.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829.

Moser, G., B. Tier, R.E. Crump, M.S. Khatkar, and H.W. Raadsma. 2009. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. Genet Sel Evol 41: 56.

Muir, W.M. 2007. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. J Anim Breed Genet 124: 342–355.

Munoz, P.R., M.F.R. Resende, D.A. Huber, T. Quesada, M.D. V. Resende, D.B. Neale, et al. 2014. Genomic Relationship Matrix for Correcting Pedigree Errors in Breeding Populations: Impact on Genetic Parameters and Genomic Selection Accuracy. Crop Sci 53: 1115–1123.

Murray, G.M., and J.P. Brennan. 2009. Estimating disease losses to the Australian wheat industry. Australas Plant Pathol 38(6): 558–570.

Newcomb, M., P.D. Olivera, M.N. Rouse, L.J. Szabo, J. Johnson, S. Gale, et al. 2016. Kenyan Isolates of Puccinia graminis f. sp. tritici from 2008 to 2014 : Virulence to SrTmp in the Ug99 Race Group and Implications for Breeding Programs. Phytopathology 106(7): 729–736.

O'Driscoll, A., S. Kildea, F. Doohan, J. Spink, and E. Mullins. 2014. The wheat-Septoria conflict: A new front opening up? Trends Plant Sci 19(9): 602–610.

Ornella, L., S. Singh, P. Perez, J. Burgueño, R. Singh, E. Tapia, et al. 2012. Genomic Prediction of Genetic Values for Resistance to Wheat Rusts. Plant Genome 5: 136–148.

Orton, E.S., S. Deller, and J.K.M. Brown. 2011. *Mycosphaerella graminicola*: From genomics to disease control. Mol Plant Pathol 12(5): 413–424.

Park, T., and G. Casella. 2008. The Bayesian Lasso. J Am Stat Assoc 103(482): 681–686.

Perez-Rodriguez, P., D. Gianola, J.M. Gonzalez-Camacho, J. Crossa, Y. Manes, and S. Dreisigacker. 2013. Comparison Between Linear and Non-parametric Regression Models for Genome-Enabled Prediction in Wheat. G3 Genes|Genomes|Genetics 2(12): 1595–1605.

Pérez, P., and G. de Los Campos. 2014. Genome-Wide Regression and Prediction with the BGLR Statistical Package. Genetics 198: 483–495.

Perez, P., G. de los Campos, J. Crossa, and D. Gianola. 2010. Genomic-Enabled Prediction Based on Molecular Markers and Pedigree Using the Bayesian Linear Regression Package in R. Plant Genome 3: 106–116.

Peterson, R.F., A.B. Campbell, and A.E. Hannah. 1948. A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. Can J Res 26c(5): 496–500.

Piepho, H.P. 2009. Ridge Regression and Extensions for Genomewide Selection in Maize. Crop Sci 49: 1165–1176.

Piyasatian, N., R.L. Fernando, and J.C.M. Dekkers. 2007. Genomic selection for marker-assisted

improvement in line crosses. Theor Appl Genet 115: 665–674.

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7(2): e32253.

Polley, R.W., and M.R. Thomas. 1991. Surveys of diseases of winter wheat in England and Wales, 1976–1988. Ann Appl Biol 119(1): 1–20.

Resende, M.F.R., P. Munoz, M.D. V. Resende, D.J. Garrick, R.L. Fernando, J.M. Davis, et al. 2012. Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (*Pinus taeda* L.). Genetics 190: 1503–1510.

Roelfs, A.P., R.P. Singh, and E.E. Saari. 1992. Rust Diseases of Wheat: Concepts and methods of disease management. CIMMYT, Mexico, D.F.

de Roos, A.P.W., B.J. Hayes, and M.E. Goddard. 2009. Reliability of Genomic Predictions Across Multiple Populations. Genetics 183: 1545–1553.

Rouse, M.N., L.E. Talbert, D. Singh, and J.D. Sherman. 2014. Complementary epistasis involving Sr12 explains adult plant resistance to stem rust in Thatcher wheat (Triticum aestivum L.). Theor Appl Genet 127(7): 1549–1559.

Rutkoski, J., J. Benson, Y. Jia, G. Brown-Guedira, J.-L. Jannink, and M. Sorrells. 2012. Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. Plant Genome J 5(2): 51.

Rutkoski, J.E., E.L. Heffner, and M.E. Sorrells. 2011. Genomic selection for durable stem rust resistance in wheat. Euphytica 179: 161–173.

Rutkoski, J.E., J.A. Poland, R.P. Singh, J. Huerta-Espino, S. Bhavani, H. Barbier, et al. 2014. Genomic Selection for Quantitative Adult Plant Stem Rust Resistance in Wheat. Plant

Genome 7(3).

Saari, E.E., and J. Prescott. 1975. A scale for appraising the foliar intensity of wheat diseases. Plant Dis Report 59(5): 376–381.

Sansaloni, C., C. Petroli, D. Jaccoud, J. Carling, F. Detering, D. Grattapaglia, et al. 2011. Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. BMC Proc 5: P54.

Shabeer, A., and W.W. Bockus. 1988. Tan spot effects on yield and yield components relative to growth stage in winter wheat. Plant Dis 72: 599–602.

Shi, G., T.L. Friesen, J. Saini, S.S. Xu, J.B. Rasmussen, and J.D. Faris. 2015. The Wheat Gene *Snn7* Confers Susceptibility on Recognition of the *Parastagonospora nodorum* Necrotrophic Effector SnTox7. Plant Genome 8(2).

Singh, P.K., E. Duveiller, and R.P. Singh. 2011. Evaluation of CIMMYT germplasm for resistance to leaf spotting diseases of wheat. Czech J Genet Plant Breed 47: S102–S108.

Singh, R.P., D.P. Hodson, Y. Jin, E.S. Lagudah, M.A. Ayliffe, S. Bhavani, et al. 2015. Emergence and Spread of New Races of Wheat Stem Rust Fungus: Continued Threat to Food Security and Prospects of Genetic Control. Phytopathology: PHYTO01150030FI.

Singh, P.K., M. Mergoum, T.B. Adhikari, S.F. Kianian, and E.M. Elias. 2006. Chromosomal location of genes for seedling resistance to tan spot and Stagonospora nodorum blotch in tetraploid wheat. Euphytica 155: 27–34.

Spindel, J., H. Begum, D. Akdemir, P. Virk, B. Collard, E. Redona, et al. 2015. Genomic Selection and Association Mapping in Rice (*Oryza sativa*): Effect of Trait Genetic Architecture, Training Population Composition, Marker Number and Statistical Model on Accuracy of Rice

Genomic Selection in Elite, Tropical Rice Breeding Line. PLoS Genet 11(2): e1004982.

Tibshirani, R. 1996. Regression Selection and Shrinkage via the Lasso. J R Stat Soc B 58: 267–288.

Toosi, A., R.L. Fernando, and J.C.M. Dekkers. 2010. Genomic selection in admixed and crossbred populations. J Anim Sci 88(1): 32–46.

Torriani, S.F.F., J.P.E. Melichar, C. Mills, N. Pain, H. Sierotzki, and M. Courbot. 2015. *Zymoseptoria tritici*: A major threat to wheat production, integrated approaches to control. Fungal Genet Biol 79: 8–12.

Vanderplank, J.E. 1963. Plant Diseases: Epidemics and Control. Academic Press, NY.

VanRaden, P.M. 2008. Efficient Methods to Compute Genomic Predictions. J Dairy Sci 91(11): 4414–4423.

Villanueva, B., R. Pong-Wong, J. Fernández, and M.A. Toro. 2005. Benefits from marker-assisted selection under an additive polygenic genetic model. J Anim Sci 83: 1747–1752.

van der Werf, J.H.J. 2009. Potential benefit of genomic selection in sheep. Proc Assoc Adv Anim Breed Genet 18: 38–41.

Whittaker, J.C., R. Thompson, and M.C. Denham. 2000. Marker-assisted selection using ridge regression. Genet Res 75: 249–252.

Wolc, A., J. Arango, P. Settar, J.E. Fulton, N.P. O'Sullivan, R. Preisinger, et al. 2011a. Persistence of accuracy of genomic estimated breeding values over generations in layer chickens. Genet Sel Evol 43(23).

Wolc, A., C. Stricker, J. Arango, P. Settar, J.E. Fulton, N.P. O'Sullivan, et al. 2011b. Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. Genet Sel Evol 43(1): 5.

De Wolf, E.D., R.J. Effertz, S. Ali, and L.J. Francl. 1998. Vistas of tan spot research. Can J Plant Pathol 20(4): 349–370.

Wong, C.K., and R. Bernardo. 2008. Genomewide selection in oil palm: Increasing selection gain per unit time and cost with small populations. Theor Appl Genet 116(6): 815–824.

Xu, S. 2003. Estimating polygenic effects using markers of the entire genome. Genetics 163: 789–801.

Yang, J., B. Benyamin, B.P. McEvoy, S. Gordon, A.K. Henders, D.R. Nyholt, et al. 2010. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42(7): 565–569.

Yu, L.X., A. Lorenz, J. Rutkoski, R.P. Singh, S. Bhavani, J. Huerta-Espino, et al. 2011. Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. Theor Appl Genet 123: 1257–1268.

Zadoks, J.C., T.T. Chang, and C.F. Konzak. 1974. A Decimal Code for the Growth Stages of Cereals. Weed Res 14(6): 415–421.

Zhang, Z., T.L. Friesen, S.S. Xu, G. Shi, Z. Liu, J.B. Rasmussen, et al. 2011. Two putatively homoeologous wheat genes mediate recognition of SnTox3 to confer effector-triggered susceptibility to *Stagonospora nodorum*. Plant J 65: 27–38.

Zhao, Y., J. Zeng, R. Fernando, and J.C. Reif. 2013. Genomic Prediction of Hybrid Wheat Performance. Crop Sci 53(3): 802–810.

Zhong, S., J.C.M. Dekkers, R.L. Fernando, and J.-L. Jannink. 2009. Factors Affecting Accuracy From Genomic Selection in Populations Derived From Multiple Inbred Lines: A Barley Case Study. Genetics 182(1): 355–364.

CHAPTER 3

GENOME WIDE ASSOCIATION STUDIES FOR RESISTANCE TO LEAF RUST, STRIPE RUST, STAGONOSPORA NODORUM BLOTCH AND TAN SPOT IN WHEAT REVEALS POTENTIAL CANDIDATE GENES

ABSTRACT

Leaf rust (LR), stripe rust (YR), Stagonospora nodorum blotch (SNB) and tan spot (TS) are some of the important foliar diseases in wheat (*Triticum aestivum* L.). To identify candidate resistance genes for these diseases in CIMMYT's (International Maize and Wheat Improvement Center) international bread wheat screening nurseries, we used genome-wide association studies (GWAS) in conjunction with information from the population sequencing map and Ensembl plants. Wheat entries in these nurseries were genotyped using genotyping-by-sequencing and phenotyped in replicated trials. Using a mixed linear model, we observed that seeding resistance to: LR was associated with twelve markers on chromosomes 1DS, 2AS, 2BL, 3B, 4AL, 6AS and 6AL; SNB was associated with ten markers on chromosomes 2AS, 2BL, 2DL, 3AS, 5BL, 6AS, 6AL and 7AL and TS was associated with fifteen markers on chromosomes 1AS, 2AL, 2BL, 3AS, 3AL, 3B, 6AS and 6AL. Seedling and adult plant resistance (APR) to YR were associated with several markers at the distal end of chromosome 2AS. In addition, YR APR was also associated with markers on chromosomes 2DL, 3B and 7DS. The potential candidate genes for these diseases included several resistance genes, receptor-like serine/threonine-protein kinases and defense-related enzymes. However, candidates of interest have to be further mapped, functionally characterized and validated. We also explored a segment on chromosome 2AS associated with multiple disease resistance and identified seventeen resistance genes. We conclude that identifying

candidate genes linked to significant markers in GWAS is feasible in wheat, thus creating opportunities for accelerating molecular breeding.

ABBREVIATIONS

ABC, Adenosine triphosphate binding cassette; APR, adult plant resistance; ARM, armadillo; BLAST, basic local alignment search tool; CIMMYT, International Maize and Wheat Improvement Center; ETS, effector-triggered susceptibility; ETI, effector-triggered immunity; GBS, genotyping-by-sequencing; GWAS, genome-wide association studies; IBWSN, international bread wheat screening nursery; IWGSC, International wheat genome sequencing consortium; LD, linkage disequilibrium; LR, leaf rust; MLM, mixed linear model; RNA, ribonucleic acid; NB-ARC, nucleotide binding-APAF-1 (apoptotic protease-activating factor-1), R proteins and CED-4 (*Caenorhabditis elegans* death-4 protein)); NBS-LRR, nucleotide binding site-leucine rich repeat; PAL, phenylalanine ammonia-lyase; PAMPs, pathogen-associated molecular patterns; POPSEQ, population sequencing; POX, peroxidase; PR, pathogenesis-related; PRRs, pattern recognition receptors; PTI, pathogen-associated molecular pattern-triggered immunity; QTL, quantitative trait loci; RGA, resistance gene analog; RLK, receptor-like kinase; RPM1, resistance to *Pseudomonas syringae* pv *maculicola* 1; RPP13, recognition of *Peronospora parasitica* 13; SINA, seven in absentia; SNB, Stagonospora nodorum blotch; STPK, serine/threonine-protein kinase; TASSEL, Trait Analysis by aSSociation Evolution and Linkage; TILLING, targeting induced local lesions in genomes; TS, tan spot; YR, stripe rust.

INTRODUCTION

Leaf rust (LR) caused by *Puccinia triticina* Eriks., stripe rust (YR) caused by *Puccinia striiformis* West., Stagonospora nodorum blotch (SNB) caused by *Parastagonospora nodorum* (Berk.) Quaedvlieg, Verkley and Crous and tan spot (TS) caused by *Pyrenophora tritici-repentis* (Died.) Shoemaker are some of the important foliar diseases in wheat (*Triticum aestivum* L.). Among these, LR or brown rust is the most common disease in many wheat producing areas of the world and can cause substantial yield losses (Roelfs et al., 1992; Marasas et al., 2004), due to reduced kernel number and kernel weight. While the early onset of disease can cause yield losses greater than 50%, losses from 7 to 30% depending on the developmental stage are common (Huerta-Espino et al., 2011). Similarly, YR is a serious disease that is prevalent in the temperate regions and results in yield losses ranging from 10 to 70% (Chen, 2005). Besides these rusts, two other foliar diseases that are globally distributed and economically significant are SNB (Shipton et al., 1971; King et al., 1983; Eyal et al., 1987) and TS (Rees et al., 1982; Shabeer and Bockus, 1988; De Wolf et al., 1998). Both diseases can cause yield losses ranging from 18% to 31% under favorable conditions (Bhathal et al., 2003) and are serious constraints to wheat production in Western Australia (Murray and Brennan, 2009). While fungicides and agronomic practices are available for the management of these diseases, the deployment of resistant cultivars is considered to be the most economical and effective strategy.

Plant resistance mechanisms against pathogens are complex. In the first line of defense, conserved molecular signatures of pathogens known as pathogen (or microbe)-associated molecular patterns (PAMPs) are recognized by plant pattern recognition receptors (PRRs) that activate the basal resistance or PAMP-triggered immunity (PTI). Successful pathogens, however, suppress PTI through secreting virulent effector proteins. These effectors activate the second line

of defense known as effector-triggered immunity (ETI) mediated by specific disease resistance (*R*) genes (Jones and Dangl, 2006). In a typical gene-for-gene interaction between a biotrophic pathogen and a plant, the effectors produced by avirulent (*Avr*) genes in the pathogen are recognized by the corresponding *R*-genes in the plant (Flor, 1956) that predominantly encode the nucleotide binding site-leucine rich repeat (NB-LRR) class of proteins (Hammond-Kosack and Jones, 1997). Upon this recognition, a hypersensitive response is initiated and leads to localized programmed cell death preventing further colonization by the pathogen (Greenberg and Yao, 2004). However, selection pressure on the pathogen imposed by large area monoculture and/or long-term deployment of varieties with single *R*-genes leads to strong selection of mutants with virulence. When the frequency of the pathogen population with virulent mutations increases, it results in the breakdown of resistance genes (McDonald and Linde, 2002). This has shifted the breeding focus from race-specific/qualitative resistance conditioned by large effect, single *R*-genes to race non-specific/quantitative resistance. Quantitative resistance is generally conditioned by many genes of small effect leading to a preferred mechanism to achieve durability (Johnson, 1984). In this type of resistance, the spread of the disease is delayed and is only expressed in adult plants (adult plant resistance, APR) in contrast to *R*-gene resistance that is usually expressed in both seedling and adult plant stages (all stage resistance). To date, more than 74 LR resistance (*Lr*) and 76 YR resistance (*Yr*) genes have been identified and most of them are race-specific except for *Lr34/Yr18/Sr57*, *Lr46/Yr29/Sr58*, *Lr67/Yr46/Sr55*, *Lr68* and *Yr36* (McIntosh et al., 2016). Combinations of *R*-genes with APR genes are expected to provide good levels of durable rust resistance (Kolmer et al., 2009; Ellis et al., 2014).

The interaction of wheat with necrotrophic fungi, *P. nodorum* and *P. tritici repentis* does not follow the gene-for-gene model. These pathogens secrete necrotrophic effectors (also known

as host-selective toxins) (Friesen et al., 2008) that interact with a corresponding host sensitivity gene and result in a compatible susceptible interaction. This is referred to as effector-triggered susceptibility (ETS) and the interaction is described as an inverse gene-for-gene model (Friesen et al., 2007). Since, susceptible cultivars could rapidly select for pathogen populations carrying the necrotrophic effectors, breeding efforts focus on eliminating the known susceptibility genes. For SNB, several interactions between necrotrophic effectors and the host genes have been identified which include: SnTox1-*Snn1* (Liu et al., 2004), SnToxA-*Tsn1* (Friesen et al., 2006; Liu et al., 2006), SnTox2-*Snn2* (Friesen et al., 2007), SnTox3-*Snn3* (Friesen et al., 2008b), SnTox3-*Snn3-B1* and SnTox3-*Snn3-D1* (Zhang et al., 2011), SnTox4-*Snn4* (Abeysekara et al., 2009), SnTox5-*Snn5* (Friesen et al., 2012), SnTox6-*Snn6* (Gao et al., 2015) and SnTox7-*Snn7* (Shi et al., 2015). For TS, six resistance genes *Tsr1/tsn1* (Faris et al., 1996)*, Tsr2/tsn2* (Singh et al., 2006a)*, Tsr3/tsn3* (Tadesse et al., 2006a)*, Tsr4/tsn4* (Tadesse et al., 2006b)*, Tsr5/tsn5* (Singh et al., 2008) and *Tsr6/tsc2* (Friesen and Faris, 2004) have been identified.

Genomics-assisted breeding for disease resistance typically involves gene identification, isolation, cloning, functional characterization to elucidate the genetic mechanism of resistance, validation and deployment. Resistance genes can be identified by either linkage mapping or genome-wide association studies (GWAS) that is based on linkage disequilibrium (LD) between a marker and the causal polymorphism. GWAS provides a much finer resolution than linkage mapping because it accounts for greater allelic diversity at a given locus and exploits the ancestral recombination events that have occurred in an existing diversity panel at the population level (Yu and Buckler, 2006). It has been successfully used to dissect several complex traits in wheat (Breseghello and Sorrells, 2006; Crossa et al., 2007; Yu et al., 2011; Juliana et al., 2015). However, several novel quantitative trait loci (QTL) identified in GWAS studies in wheat have not been

validated and functionally characterized which have limited their use in breeding programs. In addition, it is also likely that some of the novel GWAS hits may just be associated with intergenic regions and not genes *per se*. Hence, identifying the potential candidate genes linked to significant markers is important as it can provide better insights into results from GWAS. Although this was not possible with the available genetic maps in wheat, the availability of the population sequencing (POPSEQ) reference map (Chapman et al., 2015) that bridges the genetic and physical maps in wheat has made it feasible. The POPSEQ map was developed by whole-genome shotgun sequencing of wheat cultivars, 'Synthetic W7984', 'Opata' and their recombinant progenies followed by anchoring of the contigs in an ultra-dense genetic map. The POPSEQ data and the chromosome survey sequence assemblies of *T. aestivum* cv. Chinese Spring (International wheat genome sequencing consortium (IWGSC), 2014) available at Ensembl plants (Bolser et al., 2016) (http://archive.plants.Ensembl.org/Triticum_aestivum/Info/Index) provide an excellent platform for identifying genes linked to the significant markers with known physical positions in the genome. Hence, our objective was to conduct a GWAS for seedling resistance to LR, SNB and TS and explore the genes linked to the markers using Ensembl plants.

MATERIALS AND METHODS

Population

The 45th and 46th international bread wheat screening nursery (IBWSN) entries comprising 333 lines and 313 lines, respectively were used for this study. The selected bulk breeding scheme was used to develop these lines that were evaluated in cooperating locations globally. Being new advanced breeding lines from CIMMYT's (International Maize and Wheat Improvement Center)

bread wheat breeding program, they are expected to have effective and novel resistance genes which makes them ideal for association mapping.

Phenotypic evaluations for leaf rust, stripe rust, Stagonospora nodorum blotch and tan spot

Seedling evaluations for LR, SNB and TS were conducted in CIMMYT's greenhouses at El Batan, Mexico for the 45th IBWSN entries.

For LR, freshly collected urediniospores (race MBJ/SP) were suspended in light mineral oil, Soltrol (Phillips 66 Co., Bartlesville, OK, USA) and inoculation was done at the two-leaf stage. The plants were placed in a dew chamber overnight and then transferred to the greenhouse where the minimum, maximum, and average temperatures were 16.1°C, 30.0°C and 20.3°C, respectively. The 0 to 4 scale described in Roelfs et al. (1992) was used to evaluate the seedling infection types at 10 days post-inoculation. The scores were linearized to a 0-9 scale as follows: ; = 0, 0 = 0, 1- = 1, 1=2, 1+ = 3, 2- = 4, 2 = 5, 2+ = 6, 3- = 7, 3 = 8, 3+ = 9 and 4 = 9.

For SNB, the inoculum was prepared as described in Singh et al. (2006). The isolate *Sn*4 that produces SnToxA, SnTox1, SnTox2 and SnTox3 (Liu et al., 2009; Faris et al., 2011; Crook et al., 2012) was used at a concentration of 1 x 10^6 spores/ml. Four seedlings were used to represent each entry and checks Erik, Glenlea, 6B-662 and 6B-365 were planted every 20 rows. Seven days post-inoculation, the second leaf of each seedling was scored using the 1 to 5 lesion rating scale (Feng et al. 2004).

For TS, the isolate Ptr1 (Race 1) that produces PtrToxA and PtrToxC (Singh et al., 2009) was used. Inoculum preparation was done as described in Singh et al. (2011) and the concentration was adjusted to 4000 conidia/ml for both seedling and field inoculation. Seedling inoculation was done similar to that for *P. nodorum* and the same checks were used. Seedling response was

evaluated seven days post inoculation on a 1 to 5 lesion rating scale developed by Lamari and Bernier (1989a). Two replications were scored for LR and six replications were scored for SNB and TS.

Seedling and APR to YR were evaluated for the 46th IBWSN entries. While seedling evaluation was conducted in CIMMYT's greenhouses at El Batan, Mexico, APR evaluations were performed at Toluca, Mexico during the 2011 and 2013 crop seasons, at Quito, Ecuador in 2012 season and at the Kenya Agricultural and Livestock Research Organization, Njoro during the 2011 main season. For seedling evaluation, inoculum preparation and inoculation were similar to that of LR and the *P. striiformis* race, Mex96.11 was used. The seedlings were incubated in a dew chamber in the dark for 48 hours at 7°C and then transferred to the greenhouse where the minimum, maximum, and average temperatures were 6.3°C, 30.9°C and 17.3°C, respectively. The YR infection types were recorded at 14 days post-inoculation using a 0 to 9 scale as described by McNeal et al. (1971). For YR APR evaluation, the lines were sown in 0.7-m long paired rows on top of 30-cm-wide raised beds. The spreaders consisted of a mixture of six susceptible wheat lines derived from an Avocet/Attila cross. The 4-week old spreaders and hills were inoculated three times, at three to four day intervals with mixed *Pst* isolates, Mex96.11 and Mex08.13. While Mex96.11 is virulent to *Yr27* and avirulent to *Yr31*, it is the reverse for Mex08.13. Evaluations were conducted at three time points between early and late dough stages. The first evaluation was done when the severity of susceptible check, Avocet reached 80% followed by two more evaluations at weekly intervals. The modified Cobb Scale (Peterson et al. 1948) was used to score rust severity by determining the percentage of infected tissue (0-100%).

All the phenotyping data were transformed using the boxcox transformation (Box and Cox, 1964).

Genotyping and linkage disequilibrium analysis

Genome-wide markers were obtained for the lines using genotyping-by-sequencing (GBS) as described by Poland et al. (2012). Markers with missing data greater than 50% and minor allele frequency less than 10% were filtered, which resulted in in 3510 and 8072 markers with known positions for the 45th and 46th IBWSN respectively. Marker missing data was imputed using the expectation-maximization algorithm implemented in the rrBLUP software package (Endelman, 2011). After filtering the lines for missing data greater than 50%, we obtained 267 lines and 305 lines in the 45th and 46th IBWSN, respectively. The pairwise LD between the markers based on their correlations (R^2) was calculated using the 'R' statistical program.

Genome wide association mapping

Genome-wide association mapping employed the mixed linear model (MLM) (Yu et al., 2006) in TASSEL (Trait Analysis by aSSociation Evolution and Linkage) (Bradbury et al., 2007), version 5.2.24. As population structure can result in spurious associations, it was taken into account by using the first two principal components (Price et al., 2006), calculated in TASSEL using the correlation matrix. Since, there were several sibs in both nurseries, the kinship matrix obtained using the centered identity-by-state method (Endelman and Jannink, 2012) was used as a random effect to account for the degree of relatedness between sibs. The MLM was run with the optimum level of compression and the 'population parameters previously determined' method (Zhang et al., 2010). An alpha level of 0.001 was used to declare markers to be significant. To correct for multiple testing, the step up procedure of Benjamini and Hochberg (1995) which controls the false discovery rate was used with a cut-off value of 0.2. To find the candidate genes linked to significant markers, the physical starting point of the marker preceded by the

chromosome name was taken to Ensembl and a few thousand base pairs were added before and after (eg. if the position of the marker was 944423 on chromosome 2A, we used 2A: 942423-946423). The number of base pairs added varied for each marker depending on its proximity to the genes, but only the genes that were in the same genetic position were considered. The interval was then explored for predicted genes and annotations that were available from the IWGSC were obtained. For several genes, the IWGSC annotations were not available and so we evaluated orthologous genes in related species with known predicted functions using the comparative genomics tool in Ensembl. The closest species, *Triticum urartu* (A-genome donor) and *Aegilops tauschii* (D-genome donor) were first considered and when orthologs were not available or annotated in them, more distant species including barley (*Hordeum vulgare*), Brachypodium (*Brachypodium distachyon*), rice (*Oryza sativa*), maize (*Zea mays*), foxtail millet (*Setaria italica*), thale cress (*Arabidopsis thaliana)* and banana (*Musa acuminata*) were considered. In some cases, when the genes had a less similar disease resistance ortholog (<70%) in the annotated genomes of related species in Ensembl, the sequence of the *T. aestivum* gene was taken to NCBI and the nucleotide basic local alignment search tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used. This search also included the gene predictions in different species available in GenBank, but not in Ensembl. We also looked at the domains in the *T. aestivum* gene transcripts that were available in Ensembl.

RESULTS AND DISCUSSION

Phenotyping and genotyping data analysis

In the 45th IBWSN, the mean LR seedling score was 7.0 $\pm$ 2.1 on a 0-9 scale, the mean SNB seedling score was 2.5 $\pm$ 0.7 on a 1-5 scale and the mean TS seedling score was 2.6 $\pm$ 0.8 on

a 1-5 scale. The correlation of the mean LR seedling score with the mean SNB and the mean TS seedling score was very low (-0.02 and -0.11, respectively). The correlation between the mean SNB and TS seedling scores was moderate (0.47). In the 46th IBWSN, the mean YR seedling score was 6.2 $\pm$ 2.1 on a 0-9 scale. In contrast, the mean YR severities on a 0-100% severity scale were only 5.5 $\pm$ 8.8 (Quito, 2012), 6.1 $\pm$ 6.6 (Njoro 2011), 2 $\pm$ 3.2 (Toluca, 2011) and 8.7 $\pm$ 6.5 (Toluca, 2013), despite high disease pressures leading to 100% severity for the susceptible check.

We analyzed the relative percentage of markers in each chromosome and observed that chromosome 2B had the highest percentage (~12.6%) of markers in both the nurseries followed by chromosomes 3B (~11%), 5B (~8%), 2A (~7.4%) and 7A (~7%). Chromosomes 1A, 1B, 6A, 7B, 6B, 4A and 3A had about 5% of the markers each. Chromosomes 5A (~3%) and 4B (~2.5%) had the lowest percentage of markers in the A and B genomes, respectively. Overall, the D-genome had the lowest number of markers. It ranged from 1.3% to 2.2% on chromosomes 7D, 1D, 6D, 2D and 5D in both nurseries, while, chromosomes 3D and 4D had the least number of markers (less than 1%).

Linkage disequilibrium and principal component analysis

Linkage disequilibrium estimated as the allele frequency correlations (R^2) between the GBS markers across the chromosomes was plotted against the physical distance in base pairs (bps). Similar trends of LD decay were observed in both the nurseries. Hence only the LD decay for the 45th IBWSN is shown in Figure 3.1 and that for the 46th IBWSN is shown in Figure 3.2.

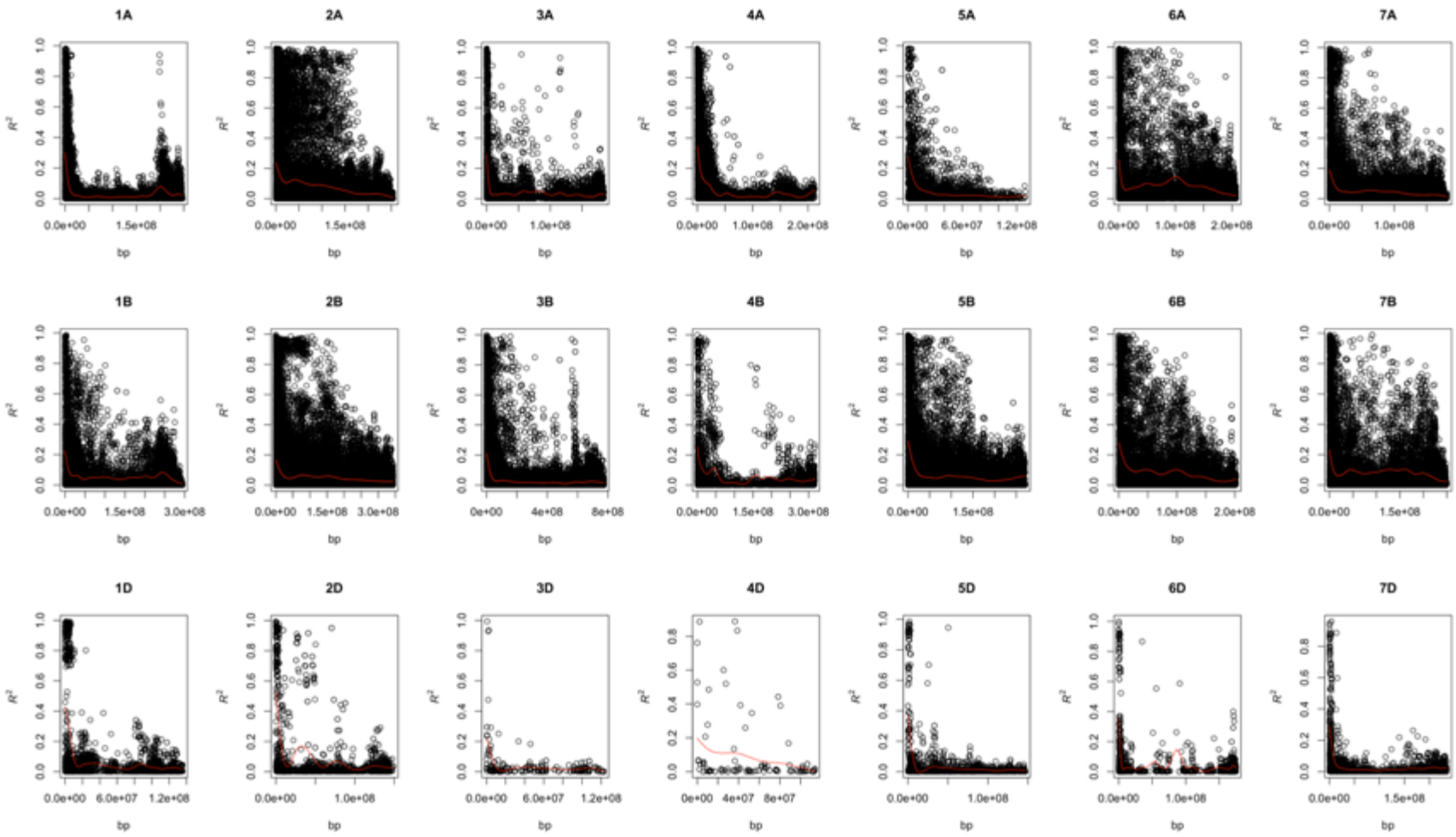

Figure 3.1: Scatter plot showing the linkage disequilibrium (LD) decay across the chromosomes. The physical distance in base pairs is plotted against the LD estimate (R^2) for pairs of markers in the 45th IBWSN.

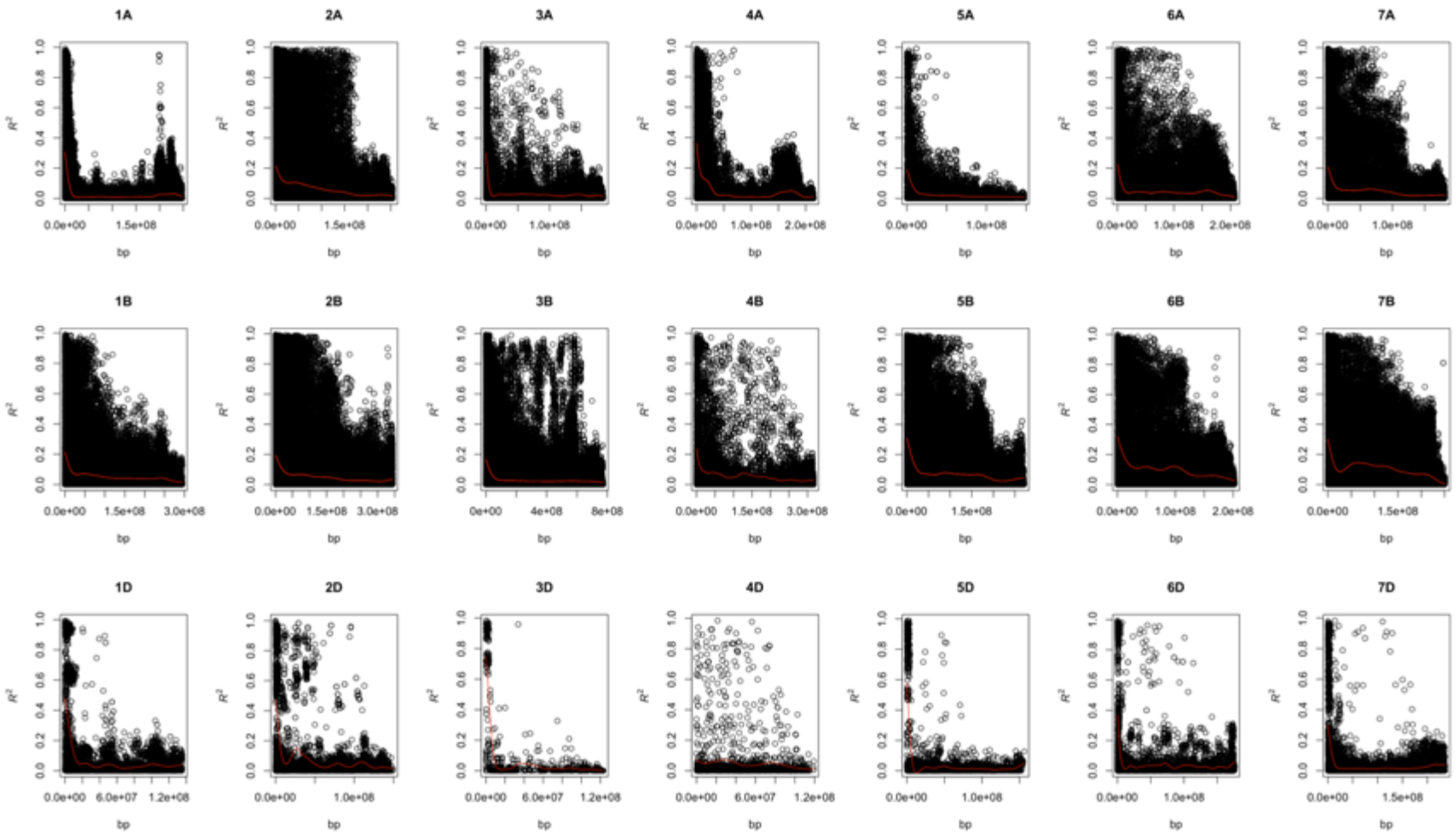

Figure 3.2: Scatter plot showing the linkage disequilibrium (LD) decay across the chromosomes. The physical distance in base pairs is plotted against the LD estimate (R^2) for pairs of markers in the 46th IBWSN.

The average extent of LD considered as the physical distance taken for the decay of R^2 to a critical value of 0.10 across the genome was approximately 5 x 10^7 bps.

Principal component analysis revealed that there was moderate population structure in both nurseries. We also identified lines with common parents and observed clear grouping of families. The lines that did not have common parents or had less than three sibs per family were classified as 'others'. In the 45th IBWSN, the first two principal components explained 9.4% and 7% of the variance, respectively (Figure 3.3).

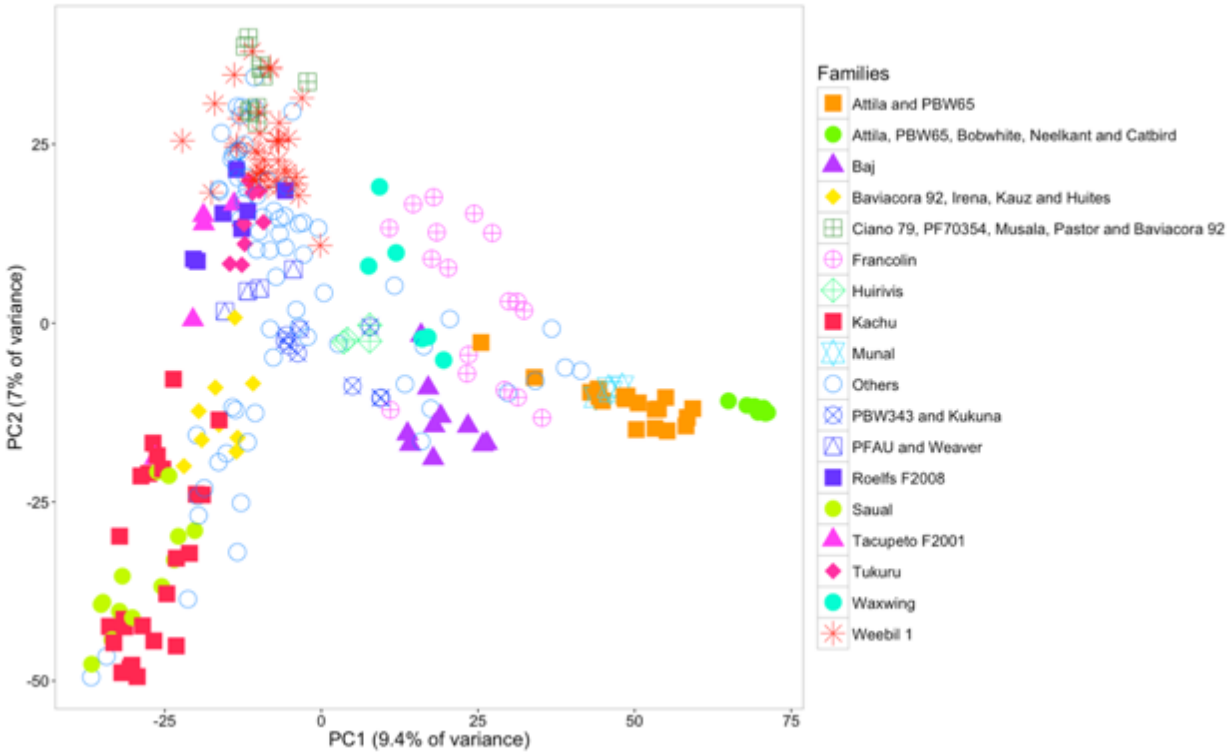

Figure 3.3: Principal component analysis and clustering of families in the 45th international bread wheat screening nursery.

Lines with 'Kachu' and 'Saual' as parents clustered together and were close to the family with 'Baviacora 92, Irena, Kauz and Huites' as parents. Lines with 'Kachu' and 'Saual' as parents

clustered together and were close to the family with 'Baviacora 92, Irena, Kauz and Huites' as parents. Lines with 'Tacupeto F2001', 'Roelfs F2008', 'Tukuru' and 'Pfau and Weaver' as parents clustered close to each other. Lines with 'PBW343 and Kukuna' and 'Huirivis' as parents clustered together. Lines with 'Ciano 79, PF70354, Musala, Pastor and Baviacora 92' and 'Weebil 1' as parents clustered together. Lines with 'Francolin', 'Baj' and 'Waxwing' formed a cluster. Lines with 'Attila and PBW65' were closely related to lines with 'Munal' in the pedigree. A family with 'Attila, PBW65, Bobwhite, Neelkant and Catbird' as parents was clearly different from others and clustered separately.

In the 46th IBWSN, the first two principal components explained 10.5% and 6% of the variance, respectively (Figure 3.4).

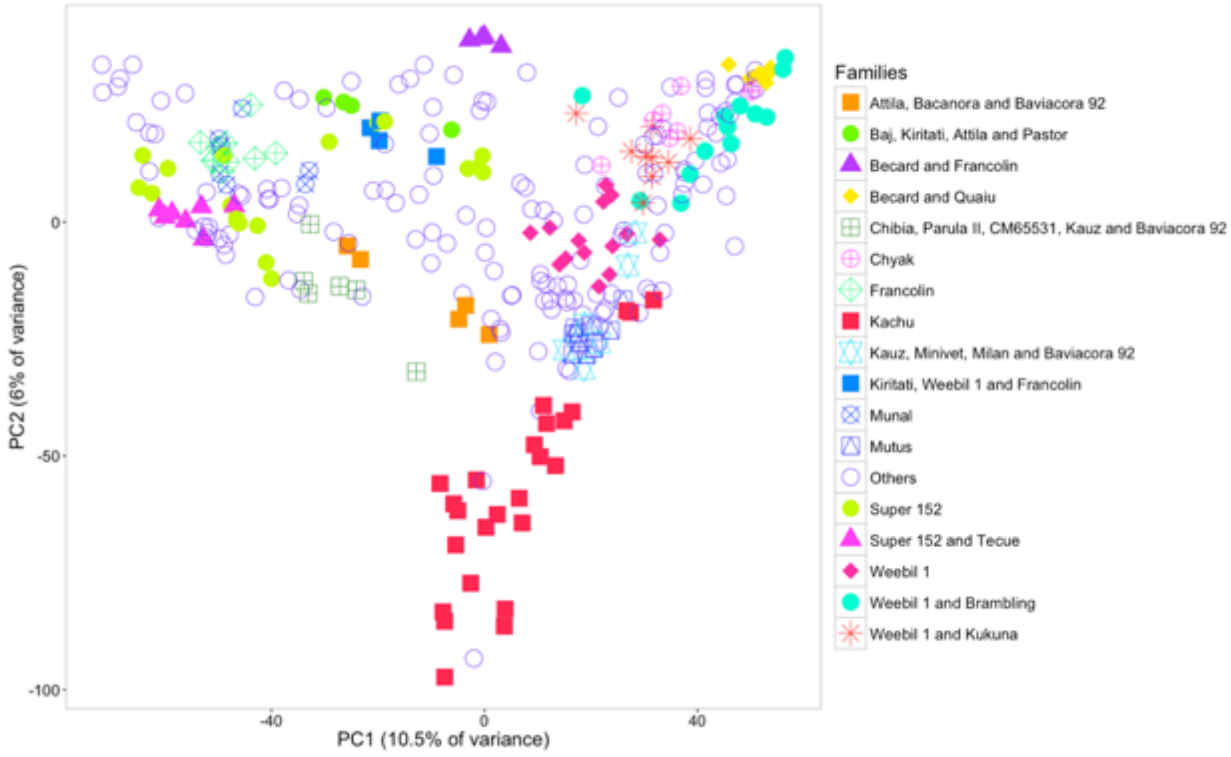

Figure 3.4: Principal component analysis and clustering of families in the 46th international bread wheat screening nursery.

Lines with 'Kachu' in the pedigree clustered together, similar to the 45th IBWSN. Lines with 'Mutus' and 'Kauz, Minivet, Milan and Baviacora 92' in the pedigree were very similar. Lines with 'Weebil 1', 'Weebil 1 and Brambling', 'Weebil 1 and Kukuna', 'Becard and Quaiu' and 'Chyak' as parents, clustered together. Sibs from a cross between 'Becard and Francolin' clustered separately. Lines with 'Munal', 'Francolin', 'Chibia, Parula II, CM65531, Kauz and Baviacora 92', 'Super 152', and 'Super 152 and Tecue' in the pedigree clustered together. Lines that had 'Baj, Kiritati, Attila and Pastor' and 'Kiritati, Weebil 1 and Francolin' in the pedigree were similar. Lines with 'Attila, Bacanora and Baviacora 92' were closer to the lines with 'Chibia, Parula II, CM65531, Kauz and Baviacora 92' in the pedigree.

Genome-wide association mapping

The markers significantly associated with LR, YR, SNB and TS, their chromosomal locations, p-values, closest *T. aestivum* gene(s), orthologous gene (only the ortholog with the highest identity is reported), the query percent identity (the percentage of the sequence in the *T. aestivum* gene that matches to the ortholog), predicted function and the domains present in the *T. aestivum* gene transcripts are reported (Tables 3.1-3.4). The adjusted p-values for the markers, R^2 values and locations of the *T. aestivum* genes are also reported (Tables S3.1-3.4). If several markers in the same genetic position were significant, only the marker with the highest significance is reported. Similarly, if several genes were in the same genetic position as the significant marker, only the adjacent gene(s) is/are reported. This is because the average LD decay was 5×10^7 base pairs and it is not feasible to report all the genes that lie within this window. Quantile-quantile plots of p-values comparing the uniform distribution of the expected $-\log_{10}$ p-value to the observed

$-\log_{10}$ p-value for different traits showed that the MLM fits the data well, except for a few datasets that had low power to detect significant associations (Supplementary Figures S3.1 and S3.2).

Markers significantly associated with seedling resistance to leaf rust

Seedling resistance to LR was associated with twelve markers: S3_6957300 (1DS), S3_1241625 (1DS), S16_199359368 (6AL), S16_5027500 (6AS), S10_147185899 (4AL), S8_40178495 (3B), S16_197872823 (6AL), S8_13948258 (3B), S8_1092429 (3B), S8_667573277 (3B), S5_344241063 (2BL) and S4_944423 (2AS) (Table 1). The two most significant markers were located on chromosome 1DS and explained 19% and 18% of the variation, respectively. This chromosome has the catalogued LR resistance genes, *Lr21* (Rowland and Kerber, 1974), *Lr42* (Cox et al., 1994a) and *Lr60* (Hiebert et al., 2008). Considering the *Lr42* gene, the marker *Xwmc432* that was tightly linked to it (Sun et al., 2010) was at 22.5cM on the wheat composite map (Somers et al., 2004). As the most significant marker in this study was at 25.4cM in the POPSEQ map, we believe it to be linked to the *Lr42* gene in that region. *Lr42* is a moderately effective race-specific resistance gene that is effective against race MBJ/SP. It originated from an *A. tauschii* introgression line, 'KS91WGRC11' (Cox et al., 1994b) and is represented as line Lr42 in CIMMYT pedigrees. This line along with CIMMYT's spring wheat line 'Quaiu' that have the *Lr42* gene (Basnet et al., 2013), were used as parents in some of the crosses and are likely the donors for resistance.

Table 3.1: Markers significantly associated with seedling resistance to leaf rust

Marker	Chromosome	Genetic position	p-value	Adjacent *T. aestivum* gene	Orthologous gene	Identity	Predicted function	Domain(s) in *T. aestivum* gene transcripts
S3_6957300	1DS	25.4	1.40E-09	Traes_1DS_3CC12E215	BRADI3G24960[†]	95	Armadillo repeat	Armadillo -type fold
				Traes_1DS_A8BD91E4A	F775_28014[‡]	100	PAL	PAL
S3_1241625	1DS	2.7	1.20E-08	Traes_1DS_3C6EAAFFD	TRIUR3_19829[§]	86	Putative disease resistance protein RGA4	NB-ARC, P-loop containing nucleoside triphosphate hydrolase
S16_199359368	6AL	118.5	3.80E-07	Traes_6AL_5A3E5FBBD	F775_22846[‡]	92	Pentatricopeptide repeat	Pentatricopeptide repeat
S16_50275005	6AS	61	8.50E-06	Traes_6AS_EB7270F83	TRIUR3_12413[§]	96	LRR receptor-like STPK	LRR
S10_147185899	4AL	29	3.30E-05	Traes_4AL_5EC714CAD	TRIUR3_02349[§]	93	Beta-glucosidase	Glycoside hydrolase, family 1
S8_40178495	3B		6.60E-05	TRAES3BF078500390CFD_g	LOC100835928[†]	85	STPK	
S16_197872823	6AL	88.2	7.30E-05	Traes_6AL_C41FC1A58			Lipid transporter	
S8_13948258	3B		7.60E-05	TRAES3BF060400070CFD_g	F775_11633[‡]	89	E3 ubiquitin-protein ligase SINA-like protein 4	
S8_1092429	3B		5.90E-04	TRAES3BF035300120CFD_g	TRIUR3_01154[§]	97	Subtilisin-like protease	
S8_667573277	3B		6.40E-04	TRAES3BF068900010CFD_g	F775_05560[‡]	96	Endoribonuclease Dicer-3a-like protein	
S5_344241063	2BL	153.8	6.50E-04	Traes_2BL_48E8EC589	F775_13446[‡]	84	Putative LRR receptor-like STPK	Concanavalin A-like lectin/glucanase domain, LRR, STPK
S4_944423	2AS	0	7.80E-04	Traes_2AS_F19BE023F	TRIUR3_09185[§]	92	Putative disease resistance protein RXW24L	Disease resistance protein, LRR, NB-ARC, P-loop containing nucleoside triphosphate hydrolase

[†]Gene from *B. distachyon*; [‡]Gene from *A. tauschii*; [§]Gene from *T. urartu*.

LRR, leucine rich repeat; NB-ARC, NB-ARC (nucleotide binding-APAF-1 (apoptotic protease-activating factor-1), R proteins and CED-4 (*Caenorhabditis elegans* death-4 protein)); PAL, Phenylalanine ammonia-lyase; RGA, resistance gene analog; SINA, seven in absentia; STPK, serine/threonine-protein kinase.

On chromosome 2AS, a marker was significant and the catalogued LR resistance genes in this chromosome are: *Lr17* from bread wheat (Dyck and Kerber, 1977)*, Lr37* from *Aegilops ventricosum* (Bariana and Mcintosh, 1993)*, Lr45* from *Secale cereale* (McIntosh et al., 1995) and *Lr65* from a Swiss spelt wheat (Mohler et al., 2012). While *Lr17*and *Lr37* are not effective against the race used, it is unlikely that *Lr45* and *Lr65* are conferring resistance in these lines given their origins. On chromosome 2BL a marker was significant and the catalogued genes in this chromosome are *Lr50* from *Triticum timopheevii subsp. armeniacum* (Brown-Guedira et al., 2003) and *Lr58* from *Aegilops triuncialis* (Kuraparthy et al., 2007). However, alien sources with these genes were not used in the crosses.

On chromosome 3B four markers were significant but their genetic positions on the POPSEQ map could not be obtained. The known LR resistance genes on this chromosome include *Lr27* from bread wheat (Singh and McIntosh, 1984) and *Lr74* that confers APR (Mcintosh et al., 2014), both of which do not confer seedling resistance to the race used. On chromosome 4AL, a marker was significant. But the catalogued LR resistance genes in this region, *Lr28* from *Triticum speltoides* (McIntosh et al., 1982) and *Lr30* from the bread wheat cultivar Terenzio (Dyck and Kerber, 1981) are unlikely to be present in this nursery as sources with these genes were not used as parents.

On chromosome 6A, two markers were significant on the long arm and one on the short arm. The known LR resistance genes on this chromosome are *Lr56* from *Aegilops sharonensis* (Marais et al., 2006), *Lr62* from *Aegilops neglecta* (Marais et al., 2009) and *Lr64* from *Triticum dicoccoides* (Mcintosh et al., 2009), all of which are located in the long arm. However, it is unlikely that any of these genes are conferring resistance in these lines, given that they were alien introgressions and were not used as parents.

Among the genes adjacent to the significant markers (Table 1), some of them could be potential candidate genes for LR resistance although they must be validated. This included disease resistance proteins, resistance gene analog 4 (RGA4) and RXW24L. The RGAs are those with sequences having homology to the conserved domains of *R*-genes like the NBS-LRR, P-loop and serine/threonine-protein kinase (STPK). While, the NBS is involved in signaling, the LRR domain mediates protein-protein interactions (Hammond-Kosack and Jones, 1997), the STPK domain phosphorylates serine and threonine and the P-loop is important for nucleotide binding (Williams et al., 2011). The RGA4 gene in particular, was found to be a constitutively active inducer of cell death (César et al., 2014). The disease resistance protein, RXW24L is a NBS-LRR gene with a P-loop, a LRR domain and a NB-ARC (nucleotide binding-APAF-1 (apoptotic protease-activating factor-1), R proteins and CED-4 (*Caenorhabditis elegans* death-4 protein) domain whose activation is known to result in cell death (van der Biezen and Jones, 1998).

In addition to the resistance genes, several STPK receptors that belong to receptor-like kinases (RLKs) were identified as potential candidates. The RLKs are a class of plant PRRs that are involved in PTI and contain an extracellular domain (LRR or lysin motif domains), a single-pass transmembrane domain and an intracellular cytosolic kinase domain (usually serine/threonine) (Shiu and Bleecker, 2001; Morris and Walker, 2003). Although LRR receptor-like STPKs are involved in several functions, the *Xa21* gene that confers resistance against bacterial blight in rice (Song et al., 1995) and flagellin-sensitive-2 gene in *Arabidopsis* that binds bacterial flagellin (Gómez-Gómez and Boller, 2000) are examples for their involvement in pathogen defense. The transcripts of several genes with LRR STPK orthologs also had a concanavalin A-like lectin/glucanase domain. Concanavalin A is a lectin that binds carbohydrates and this domain has a sandwich structure made of ß-strands found in many proteins

(http://www.ebi.ac.uk/interpro/entry/IPR013320). Some lectin RLKs are known to be involved in plant innate immunity (Bouwmeester et al., 2011; Singh et al., 2012).

Repeats belonging to the armadillo (ARM) family and the pentatricopeptide family were potential candidates. ARM repeats were initially identified in the *Drosophila* segment polarity gene, *armadillo* (Nusslein-Volhard and Wieschaus, 1980; Riggleman et al., 1989) and are a class of helical repeat proteins involved in protein interactions. The largest class of ARM repeats in *Arabidopsis* contain the U-box domain found in the E3 ubiquitin ligases, involved in the ubiquitination and targeting of proteins for proteasomal degradation (Mudgil et al., 2004). While U-box/ARM repeats have several functions, the rice *Spotted leaf11* gene encoding a U-box/ARM protein negatively regulates plant cell death and was suggested to be involved in the basal defense signaling against rice blast (Zeng et al., 2004). Pentatricopeptide repeat-containing proteins include a large gene family characterized by tandem arrays of 35 amino-acid repeats (Small and Peeters, 2000). They are ribonucleic acid (RNA)-binding proteins known to play important roles in post-transcriptional processes within the mitochondria and chloroplasts (Delannoy et al., 2007). Although they play several physiological roles, they are also known to be involved in defense against necrotrophic fungi (Laluk et al., 2011) and diverse pathogens (Park et al., 2014).

Several genes encoding enzymes like beta-glucosidase, E3 ubiquitin-protein ligases, endoribonuclease Dicer, phenylalanine ammonia-lyase (PAL) and subtilisin-like protease were also identified as potential candidates. Beta-glucosidases belong to the family 1 glycoside hydrolases that are known to activate phytoanticipins and serve as triggers of chemical defense in plants against pathogens (Nisius, 1988; Suzuki et al., 2006; Morant et al., 2008). E3 ubiquitin-protein ligases play a major role in substrate specificity and facilitate the formation of an isopeptide bond between ubiquitin and the target protein, that is subsequently degraded via the proteasome.

They are classified based on the domains they contain and some of them are known to be involved in plant defense (Zeng et al., 2004; Yang et al., 2006; Craig et al., 2009; Dielen et al., 2010). A SINA ligase, SINA3 was recently found to be involved in defense signaling and ubiquitination of a defense related transcription factor in tomato, suggesting a negative role in plant defense response (Miao et al., 2016).

Endoribonuclease Dicer-like proteins have RNase III domains that cleave double stranded RNA into 21–26 nucleotide small RNAs, that are loaded into Argonaute proteins to induce silencing of their complementary target genes (Baulcombe, 2004). Dicer-like 3, in particular is required for chromatin silencing (Xie et al., 2004). Dicer and dicer-like proteins are known to regulate plant immunity against an array of pathogens including fungi via the small RNAs processed by them (Gupta et al., 2012; Li et al., 2014; Weiberg et al., 2014). Phenylalanine ammonia lyase (EC 4.3.1.24) is a key enzyme in the phenylpropanoid pathway of higher plants involved in the production of several compounds like lignins, coumarins and flavonoids that are related to plant defense (Dixon et al., 2002; La Camera et al., 2004). Several studies have reported the induction of the PAL gene in response to fungal elicitors and its association with enhanced fungal defense (Thorpe and Hall, 1984; Edwards et al., 1985; Southerton and Deverall, 1990; Pellegrini et al., 1994; Shadle et al., 2003; Tonnessen et al., 2015). Interestingly, the wheat PAL gene had highly similar orthologs in several other plants indicating that it is conserved across species as observed by Rawal et al. (2013). Subtilisin-like proteases are serine proteases that belong to a subfamily of PR genes and have a catalytic triad of histidine, aspartate and serine residues (Siezen and Leunissen, 1997). Some of them are known to activate defense related genes and are involved in defense responses (Tornero et al., 1997; Jordá and Vera, 2000; Pearce et al., 2010).

In addition to the disease resistance genes, STPKs and enzymes, a gene encoding a lipid transporter was also a potential candidate. Lipid transport proteins transfer phospholipids between membranes (Kader, 1996) and have been classified as a PR protein family member (PR-14) (van Loon and van Strien, 1999). While they play diverse roles, they are also known to be involved in systemic resistance signaling (Maldonado et al., 2002) and inhibition of bacterial and fungal pathogens (Segura et al., 1993; Regente et al., 2005; Sarowar et al., 2009).

Markers significantly associated with seedling resistance to Stagonospora nodorum blotch

Seedling resistance to SNB was associated with ten markers: S7_3950435 (3AS), S16_339241 (6AS), S5_332672189 (2BL), S6_36849900 (2DL), S5_288066166 (2BL), S4_944423 (2AS), S19_163864204 (7AL), S4_508877 (2AS), S14_265087281 (5BL) and S16_201351017 (6AL) (Table 2). The variation explained by the significant markers ranged from 5-7%, that supports previous studies suggesting SNB resistance to be complex and quantitatively inherited (Fried and Meister, 1987; Bostwick et al., 1993; Du et al., 1999; Wicki et al., 1999). The most significant marker was on chromosome 3AS where no SNB resistance gene has been catalogued. However, Liu et al. (2004b) identified a QTL on this chromosomal arm that confers resistance to the toxins produced by the isolate Sn2000. But the relative position of the linked marker *Xksu912*(Prp), could not be obtained on the POPSEQ map. Abeysekara et al. (2009) also identified another QTL, *QSnb.fcu-3A* (6cM) conferring seedling resistance, but the isolate, 1A7a used produces SnTox4 that is not known to be produced by the isolate we used.

On chromosome 2AS, three markers were significant and no seedling resistance SNB QTL has been identified in these locations. On chromosome 2BL, two markers were significant and only a QTL (*QSnl.ihar-2B*) has been identified in this chromosomal arm using mixed isolates

(Arseniuk et al., 2004). Marker *gwm501*, that flanked the QTL was located at 94.3cM in the consensus map (Yu et al., 2014). While one of our markers was located at 99.1cM, it was not possible to determine whether they are referring to the same QTL. On chromosome 2DL, a marker was significant and the catalogued gene in this chromosome is *Snn7*. Since, Sn4 is not known to produce SnTox7 it is unlikely that this sensitivity gene was linked to the significant marker. On chromosome 5BL, a marker was significant and the cloned *Tsn1* gene (with STPK and NBS-LRR domains) is the catalogued gene in this chromosome (Faris et al., 2010). However, the likely candidate gene linked to the marker in our study was a putative LRR receptor-like STPK suggesting that it is a different gene.

On chromosome 6A, two markers that were located in the long and short arm were significant. Although the sensitivity gene, *Snn6* has been mapped to the long arm of chromosome 6A (Gao et al., 2015), it is unlikely that it was linked to the significant marker in our study as Sn4 is not known to produce SnTox6. Arseniuk et al. (2004) used mixed isolates and identified QSnl.ihar-6A on chromosome 6AL that was flanked by markers *Xgwm570* and *Xmwg934*. The marker, *Xgwm570* was located at 119.3cM in the consensus map (Yu et al., 2014). The significant marker in this study was at 123.9cM and might be identifying the same QTL. Finally, a marker was significant on chromosome 7AL. Although a QTL has been detected on this chromosomal arm (Adhikari et al., 2011), in the same location as marker, wPt-4515 that was at 52.4cM in the consensus map (Yu et al., 2014), the marker significant in this study was at 129.3cM and thus identified a different QTL.

Among the genes adjacent to the markers significantly associated with SNB (Table 2), were disease resistance protein RXW24L, receptor-like STPKs and endoribonuclease (discussed earlier), wall-associated receptor kinase, adenosine triphosphate (ATP)-binding cassette (ABC)

transporters (ABCB4 and ABCC9) and other genes encoding a non-specific lipid-transfer protein and a zinc finger CCCH domain-containing protein. Wall-associated receptor kinases are tightly bound to pectin in the cell wall and possess a cytoplasmic protein kinase domain that link and mediate signals between the plasma membrane and the cell wall (He et al., 1996; Decreux and Messiaen, 2005). WAK1, well-studied among these receptors is a PR protein that is induced by pathogen infection and salicylic acid (He et al., 1998, 1999; Brutus et al., 2010). The ABC transporters have two hydrophobic transmembrane domains that form the pathway through which substrates like sugars, amino acids, oligopeptides, inorganic ions, polysaccharides, proteins etc. cross the cell membrane and two nucleotide-binding domains that are located at the cytoplasmic side bind ATP and facilitate the transport process (Higgins, 1992). Plant ABC transporters are classified into several sub-families (ABCA - ABCH) and play diverse roles (Rea, 2007; Verrier et al., 2008). Some members of the ABCB sub-family (also known as multi-drug resistance proteins) including ABCB4 are known to be involved in auxin transport (Noh et al., 2001; Geisler et al., 2005; Cho et al., 2012). Although, there is no report showing a role for ABCB4 in plant defense, auxin transport is known to modulate plant defense response (Wang et al., 2007; Llorente et al., 2008; Kazan and Manners, 2009). The other candidate, ABCC9 transporter is a sulfonylurea receptor acting as a potassium channel regulator in humans (Bryan et al., 2007). But, its role in plants is not elucidated.

Non-specific lipid transfer proteins facilitate the transfer of a wide range of lipids and are known to be involved in plant defense against bacterial and fungal pathogens (Molina et al., 1993; García-Olmedo et al., 1995; Kristensen et al., 2000; Wang et al., 2004). Zinc finger CCCH (C and H denote cysteine and histidine, respectively) domain-containing proteins are transcriptional factors with a motif containing three cysteine residues and one histidine residue. While they play

key roles in gene expression, RNA-binding, growth and developmental processes and stress tolerance (Wang et al., 2008), one of them has been reported to confer fungal resistance in cotton (Guo et al., 2009).

Table 3.2: Markers significantly associated with seedling resistance to Stagonospora nodorum blotch

Marker	Chromosome	Genetic position	p-value	Adjacent *T. aestivum* gene	Orthologous gene	Identity	Predicted function	Domain(s) in *T. aestivum* gene transcripts
S7_3950435	3AS	9.5	2.70E-05	Traes_3AS_08B918CFD	OS01G0136400[¶]	80	STPK	STPK
S16_339241	6AS	0.9	5.70E-05	Traes_6AS_8684E560A	LOC100832443[†]	81	LRR receptor-like STPK	Concanavalin A-like lectin/glucanase domain, LRR, STPK
S5_332672189	2BL	134	1.40E-04	Traes_2BL_6B925F45E	F775_19193[‡]	93	Zinc finger CCCH domain-containing protein 32	Zinc finger, CCCH-type
S6_36849900	2DL	76.5	1.50E-04	Traes_2DL_85AFEAF70	F775_32342[‡]	80	Non-specific lipid-transfer protein	Plant lipid transfer protein/Par allergen
S5_288066166	2BL	99.1	2.30E-04	Traes_2BL_1F6C61302	GSMUA_Achr1 1G23280_001[§§]	76	ABC transporter C family member 9	ABC transporter type 1, transmembrane domain, P-loop containing nucleoside triphosphate hydrolase
S4_944423	2AS	0	4.00E-04	Traes_2AS_F19BE023F	TRIUR3_09185[§]	92	Putative disease resistance protein RXW24L	Disease resistance protein, LRR, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
S19_163864204	7AL	129.3	4.00E-04	Traes_7AL_A647529F4	F775_09042[‡]	94	Putative LRR receptor-like STPK	Protein kinase domain, STPK
S4_508877	2AS	0	4.10E-04	Traes_2AS_6A15EE669	LOC100842644[†]	85	ABC transporter B family member 4	ABC transporter type 1, transmembrane domain, P-loop containing nucleoside triphosphate hydrolase
S14_265087281	5BL	162.5	4.40E-04	Traes_5BL_D8830A0DF	F775_09968[‡]	96	Putative LRR receptor-like STPK	Protein kinase domain, STPK
	6AL	123.9	5.10E-04	Traes_6AL_671AFD8EF			Endoribonuclease	

S16_201351017				Traes_6AL_EF774ECF2	F775_02539‡	74	Wall-associated receptor kinase	Concanavalin A-like lectin/glucanase domain, Epidermal growth factor-like calcium-binding domain, STPK, Wall-associated receptor kinase galacturonan-binding domain

†Gene from *B. distachyon*; ‡Gene from *A. tauschii*; §Gene from *T. urartu*; ¶Gene from *O. sativa Japonica*; §§Gene from *M. acuminata*.

ABC, Adenosine triphosphate binding cassette; LRR, leucine rich repeat; STPK, serine/threonine-protein kinase.

Markers significantly associated with seedling resistance to tan spot

Seedling resistance to TS was associated with fifteen markers: S1_3589926 (1AS), S7_182028651 (3AL), S4_239686345 (2AL), S8_12198705 (3B), S8_13415415 (3B), S16_4196814 (6AS), S8_7801088 (3B), S8_1092429 (3B), S7_4804454 (3AS), S16_9839264 (6AS), S16_191519837 (6AL), S1_2331617 (1AS), S7_4563676 (3AS), S5_281016023 (2BL) and S1_2584791 (1AS) (Table 3). The most significant marker (explained 10% of the variation) and two other significant markers were located on chromosome 1AS (27.2cm and 28cM), where the catalogued gene is *Tsc1* (Effertz et al., 2002). Marker *Xgwm136* that was 4.7cM distal to *Tsc1* was at 11cM in a consensus map (Yu et al., 2014). But it was not possible to determine if the significant markers are linked to this gene. Chromosomes 2AL and 2BL, had a significant marker but no TS resistance gene has been reported in these chromosomes.

On chromosome 3AS, two significant markers were located in the same genetic position (12.6cM). *Tsr4*, the catalogued gene in this chromosome was 14.9cM away from the marker *Xgwm2* (Tadesse et al., 2010). *Xgwm2* was at 37cM on the wheat composite map (Somers et al., 2004) which puts *Tsr4* at about 52cM. Hence, the significant markers in this study are unlikely to be in the location of the *Tsr4* gene. On the long arm of chromosome 3A, a marker was significant, but no resistance gene has been reported in that location. On chromosome 3B, four markers were significant, but the position of only one marker could be obtained on the POPSEQ map. The known genes on chromosome 3BL are *Tsr2/tsn2* that confers resistance to the necrosis induced by a race 3 isolate (Singh et al., 2006a) and *Tsr5/tsn5* that confers resistance to the necrosis induced by a race 5 isolate (Singh et al., 2008). Faris and Friesen (2005) also identified a race non-specific QTL (QTs.fcu-3BL) on chromosome 3BL. While, the marker whose position was known (6.8cM) was not in the position of any of the known genes or QTL, it was not possible to determine if the other

markers coincide with them. Finally, three markers on chromosome 6A were significant and no TS resistance gene has been identified in this chromosome.

Among the genes adjacent to the markers significantly associated with TS (Table 3), were genes encoding disease resistance proteins, RGA4 (discussed earlier), RPM1 (resistance to *Pseudomonas syringae pv maculicola* 1), disease resistance response protein 206 and RPP13 (recognition of *Peronospora parasitica* 13). RPM1 is a coiled coil-NBS-LRR disease resistance protein that functions at the plasma membrane by interacting with another plasma membrane localized protein called RPM1-interacting protein 4 and mediates ETI to *P. syringae* (Debener et al., 1991; Mackey et al., 2002). The cloned wheat leaf rust resistance gene, *Lr10* has been reported to be similar to the RPM1 gene (Feuillet et al., 2003). The disease resistance response protein 206 is known to be involved in non-host disease resistance response (Wang et al., 1999; Wang and Fristensky, 2001; Choi et al., 2004). It is related to the dirigent protein that is suggested to play a role in conifer defense by lignan and lignin formation (Ralph et al., 2006). RPP13 is a leucine zipper NBS-LRR gene from *Arabidopsis* conferring resistance to several different isolates of *Perenospora parasitica* causing downy mildew (Bittner-Eddy et al., 2000; Bittner-Eddy and Beynon, 2001).

In addition to the disease resistance genes, receptor-like STPKs, wall-associated receptor kinases, enzymes like PAL and subtilisin, Arm repeat protein (all of which have been discussed earlier), cysteine-rich receptors, glutamate receptors, ABCC15 transporter, ABCD1 transporter, peroxidase, Bowman-Birk trypsin inhibitor and hydroxyproline-rich glycoproteins were also identified as potential candidates. Cysteine rich receptors are characterized by cysteine residues and repeats of the domain of unknown function 26 on the extracellular domain (Chen, 2001). They are known to be induced by pathogen infection and regulate basal plant defense (Acharya et al.,

2007; Ederli et al., 2011; Yeh et al., 2015). Glutamate receptors are non-selective cation channels that were initially proposed to play a role in light-signal transduction and then found to be associated with Ca^{2+} influx and metabolic signaling (Lam et al., 1998; Kim et al., 2001; Davenport, 2002). They are also known to be involved in plant defense signaling (Kang et al., 2006; Vatsa et al., 2011; Li et al., 2013).

Some members of the ABCC transporters are known to be involved in biotic stresses (Wanke and Üner Kolukisaoglu, 2010), but the function of the ABCC15 transporter in plants is not understood. The ABCD family members are peroxisomal transporters that are involved in the transfer of fatty acids (Theodoulou et al., 2006). However, the functions of plant ABCD transporters except the COMATOSE gene, that controls the switch from seed dormancy to germination (Footitt et al., 2002) are unknown. The enzyme peroxidase (POX, EC 1.11.1.7) is an important component of PTI and its activity leads to the production of reactive oxygen species in response to pathogen attack (Kawano, 2003; Daudi et al., 2012; Mammarella et al., 2015). A specific type of POX that catalyzes lignification by causing cell wall reinforcement and enhanced resistance against multiple pathogens is classified as the PR protein, PR-9 (van Loon and van Strien, 1999). Increase in POX activity in response to fungal infection has been reported in several studies (Seevers and Daly, 1970; Thorpe and Hall, 1984; Southerton and Deverall, 1990). Bowman-Birk type trypsin inhibitor is a serine protease inhibitor (Bowman, 1946; Birk et al., 1963) with antifungal activity (Terras et al., 1993; Ye et al., 2001; Qu et al., 2003; Kuhar et al., 2013). Hydroxyproline-rich glycoproteins are integral components of the primary cell wall of plants (Lamport and Northcote, 1960) that accumulate in defense response to various pathogens (Showalter et al., 1985; Corbin et al., 1987; Shailasree et al., 2004).

Table 3.3: Markers significantly associated with seedling resistance to tan spot

Marker	Chromosome	Genetic position	p-value	Adjacent *T. aestivum* gene	Orthologous gene	Identity	Predicted function	Domain(s) in *T. aestivum* gene transcript
S1_3589926	1AS	28	2.40E-05	Traes_1AS_BF353B963	LOC100824961[†]	79	Putative disease resistance protein RPM1	NBS-LRR type resistance protein
				Traes_1AS_F098402B4	TRIUR3_05134[§]	98	Bowman-Birk type trypsin inhibitor	Proteinase inhibitor I12, Bowman-Birk
S7_182028651	3AL	197.4	3.60E-04	Traes_3AL_91749D67D	F775_07165[‡]	94	Putative disease resistance protein RGA4	Powdery mildew resistance protein, LRR, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
S4_239686345	2AL	113.9	3.60E-04	Traes_2AL_34A3B95BE	GSMUA_Achr11G12630_001[§§]	42	Putative disease resistance response protein 206	Plant disease resistance response protein
					LOC101765197[¶¶]	73	Dirigent protein 1-like	
				Traes_2AL_97FC5264A	TRIUR3_01787[§]	96	Wall-associated receptor kinase-like	Protein kinase domain, STPK
S8_12198705	3B		3.80E-04	TRAES 3BF270500020CFD_g	OS01G0115750[¶]	76	STPK	
S8_13415415	3B		4.30E-04	TRAES 3BF060400190CFD_g	OS02G0626100[¶]	80	Phenylalanine ammonia-lyase	
S16_4196814	6AS	13.1	4.90E-04	Traes_6AS_3A682BA20	OS02G0106900[¶]	74	STPK	Protein kinase domain, STPK
S8_7801088	3B	6.82	5.05E-04	TRAES 3BF060200040CFD_g	BRADI3G16550[†]	87	Hydroxyproline-rich glycoprotein family	
				TRAES 3BF060200010CFD_g			ARM repeat superfamily protein	
S8_1092429	3B		8.60E-04	TRAES 3BF035300120CFD_g	TRIUR3_01154[§]	97	Subtilisin-like protease	
S7_4804454	3AS	12.6	9.00E-04	Traes_3AS_769E90DDD	TRIUR3_11178[§]	90	Putative cysteine-rich receptor-like protein kinase	Protein kinase domain, STPK
S16_9839264	6AS	40.2	9.00E-04	Traes_6AS_30C919428	F775_10287[‡]	92	Glutamate receptor	Ionotropic glutamate receptor

S16_191519837	6AL	67.9	1.40E-03	Traes_6AL_4815187D4	LOC100839119[†]	78	LRR receptor-like STPK	Concanavalin A-like lectin/glucanase domain, LRR, STPK
S1_2331617	1AS	27.2	2.30E-03	Traes_1AS_AAB89883E1	TRIUR3_12921[§]	99	Disease resistance protein RPM1	LRR, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
				Traes_1AS_459048879	F775_18040[‡]	85	Disease resistance protein RPP13	
				Traes_1AS_3BE2A2127	F775_01616[‡]	76	Putative disease resistance protein	LRR domain, L domain-like
S7_4563676	3AS	12.6	2.30E-03	Traes_3AS_817ECEF75	F775_18635[‡]	78	Wall-associated receptor kinase 1	STPK
S5_281016023	2BL	97.4	3.30E-03	Traes_2BL_7C2F474DE	TRIUR3_04133[§]	96	Peroxidase	Plant peroxidase
				Traes_2BL_D055B271C	LOC100825682[†]	94	Putative ABC transporter C family member 15	ABC transporter type 1, transmembrane domain
S1_2584791	1AS	28	4.50E-03	Traes_1AS_C8A8A4118	LOC100823561[†]	90	ABC transporter D family member 1	
				Traes_1AS_B716E0B0E	LOC100831913[†]	79	Disease resistance protein RPP13	LRR
				Traes_1AS_8D33AB43B	TRIUR3_19998[§]	91	ABC transporter D family member 1	ABC transporter type 1, transmembrane domain, P-loop containing nucleoside triphosphate hydrolase, Peroxisomal fatty acyl CoA transporter
				Traes_1AS_C7A8188D1	LOC100830206[†]	80	Disease resistance protein RPM1	Disease resistance protein, Coiled-coils

[†]Gene from *B. distachyon*; [‡]Gene from *A. tauschii*; [§]Gene from *T. urartu*; [¶]Gene from *O. sativa Japonica*; [§§]Gene from *M. acuminata*, [¶¶]Gene from *S. italica*.

ABC, Adenosine triphosphate binding cassette; ARM, armadillo; LRR, leucine rich repeat; RGA, Resistance gene analog; RPM1, resistance to *Pseudomonas syringae pv maculicola* 1; RPP13, recognition of *Peronospora parasitica* 13; STPK, serine/threonine-protein kinase.

Markers significantly associated with seedling and adult plant resistance to stripe rust

Seedling resistance to YR was associated with markers: S4_208035, S4_508877, S4_944423, S4_5007061, S4_5287800, S4_7117805 on chromosome 2AS (Table 4). All these markers except S4_5007061 (that was only significant in Quito, 2012), were also associated with APR in all the four datasets. The most significant markers for seedling resistance and APR explained 27% and 14% of the variation. As these markers were significantly associated with both seedling and APR, they are likely to be linked to an all-stage resistance gene that might be *Yr17* or a closely linked gene. The *Yr17* gene is located between 0-4cM in the wheat composite map (Somers et al., 2004) which is also the approximate location of our markers (0 cM, 8.9 cM). The gene, *Yr17* was introgressed into the French wheat cultivar 'VPM-1' as a translocation segment from the D-genome of *Aegilops ventricosa.* Lines with Kachu, Milan and Mutus are expected to have the *Yr17* gene and they were used as parents for several crosses.

The other markers significantly associated with YR include S6_132714407 (2DL) in the Njoro 2011 dataset, S8_17773150 (3B) in the Quito 2012, S8_566227604 (3B) in the Toluca 2011 dataset and S21_4853558 (7DS) in the Toluca 2013 dataset. On chromosome 2DL, three genes: *Yr37* from *Aegilops kotschyi* (Marais et al., 2005), *Yr54* from the common spring wheat line Quaiu (Basnet et al., 2013) and *Yr55* (Mcintosh et al., 2014) have been catalogued. It is unlikely that the gene linked to the significant marker is *Yr37* because *Aegilops kotschyi* was not used in the crosses. The relative position of the *Yr55* gene to the marker could not be obtained. Considering the *Yr54* gene, it is unlikely that this marker is linked to it, although it is present in 'Quaiu' that was used as a parent in some of the crosses. This is because *Xgwm301*, the marker linked to *Yr54* was at 107cM in the wheat composite map (Somers et al., 2004) and the significant marker is at 82.4cM. On chromosome 3B, two markers were significant but their positions could not be obtained. The

catalogued rust resistance genes in this chromosome are *Yr4* from common wheat (Bansal et al., 2009)*, Yr30/Sr2* that occurs in a high frequency in CIMMYT germplasm (Singh et al., 2005)*, Yr57* (Randhawa et al., 2015) and *Yr58* (Chhetri et al., 2016). On chromosome 7DS, a marker at 3.7cM was significant and *Yr18/Lr34* is the catalogued gene in this chromosome which is present in a significant frequency in the CIMMYT germplasm. But the position of a gene-specific marker, cssfr5 for the *Lr34* gene (Lagudah et al., 2009) at 49.3cM in the consensus map (Yu et al., 2014) indicates that it is not the gene linked to the significant marker.

The genes adjacent to the markers significantly associated with YR seedling resistance include PAL, ABCB4, disease resistance protein RXW24L, disease resistance RPP13-like protein and disease resistance protein RGA3. For YR APR, in addition to these, genes encoding wall-associated receptor kinase, cysteine-rich receptor like protein kinase, LRR receptor-like STPK and Mlo-like protein were also potential candidates. All the candidates except for the Mlo-like protein have been discussed earlier. The Mlo locus in barley has recessive mutations that confer broad spectrum resistance to all known isolates of the powdery mildew fungus (*Blumeria graminis* f. sp. *hordei*) (Jørgensen, 1992) and could be a potential candidate.

Table 3.4: Markers significantly associated with seedling and adult plant resistance to stripe rust

Dataset	Marker	Chromosome	Genetic position	p-value	Adjacent *T. aestivum* gene	Orthologous gene	Identity	Predicted function	Domain(s) in *T. aestivum* gene transcript
Njoro 2011	S4_208035	2AS	0	5.50E-04	Traes_2AS_5EB59FFC0	F775_06675‡	92	PAL	Aromatic amino acid lyase
Quito 2012				4.60E-04					
Seedling				2.80E-09					
Toluca 2011				3.90E-08					
Toluca 2013				3.30E-08					
Njoro 2011	S4_508877	2AS	0	6.10E-04	Traes_2AS_6A15EE669	LOC100842644†	100	ABC transporter B family member 4-like	ABC transporter type 1, transmembrane domain, P-loop containing nucleoside triphosphate hydrolase, AAA+ ATPase domain
Quito 2012				7.70E-08					
Seedling				5.80E-14					
Toluca 2011				3.00E-07					
Toluca 2013				2.40E-10					
Njoro 2011	S4_944423	2AS	0	2.50E-05	Traes_2AS_F19BE023F	TRIUR3_09185§	92	Putative disease resistance protein RXW24L	-
Quito 2012				5.90E-07					
Seedling				7.70E-13					
Toluca 2011				2.40E-07					
Toluca 2013				6.00E-11					
Quito 2012	S4_5007061	2AS	8.9	1.90E-03	Traes_2AS_6BC67DD45	TRIUR3_16539§	97	Putative disease resistance RPP13-like protein	Leucine-rich repeat domain, L domain-like
Seedling				1.90E-07					
Njoro 2011	S4_5287800	2AS	8.9	6.50E-05	Traes_2AS_A477CDA77	TRIUR3_30356§	71	Putative disease resistance RPP13-like protein	P-loop containing nucleoside triphosphate hydrolase
Quito 2012				2.60E-07					
Seedling				6.40E-11					
Toluca 2011				1.10E-06					
Toluca 2013				3.10E-08					

Njoro 2011	S4_7117805	2AS	8.9	4.30E-05	Traes_2AS_6CE6AB560	LOC100830175[†]	79	Putative disease resistance protein RGA3	Leucine-rich repeat domain, L domain-like, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Quito 2012				3.60E-06					
Seedling				3.40E-11					
Toluca 2011				8.20E-06					
Toluca 2013				1.00E-08					
Njoro 2011	S6_132714407	2DL	82.4	1.20E-03	Traes_2DL_4B5D621C1	F775_25858[‡]	100	Wall-associated receptor kinase	Concanavalin A-like lectin/glucanase domain, STPK
Quito 2012	S8_17773150	3B		2.00E-05	TRAES3BF050800140CFD_g	TRIUR3_18467[§]	91	Cysteine-rich receptor-like protein kinase	Concanavalin A-like lectin/glucanase domain, STPK
Toluca 2011	S8_566227604	3B		9.30E-04	TRAES3BF027700080CFD_g	F775_04751[‡]	94	Putative LRR receptor-like STPK	Concanavalin A-like lectin/glucanase domain, LRR, STPK
Toluca 2013	S21_4853558	7DS	3.7	3.00E-04	Traes_7DS_600B0996B	Si017007m.g[¶¶]	70	Mlo-like protein	Mlo-related protein

[†]Gene from *B. distachyon*; [‡]Gene from *A. tauschii*; [§]Gene from *T. urartu*; [¶¶]Gene from *S. italica*.

ABC, Adenosine triphosphate binding cassette; LRR, leucine rich repeat; PAL, Phenylalanine ammonia-lyase; RGA, Resistance gene analog; RPP13, recognition of *Peronospora parasitica* 13; STPK, serine/threonine-protein kinase.

A segment in the distal end of chromosome 2AS is rich in disease resistance genes

Marker S4_944423 on chromosome 2AS, was associated with seedling resistance to LR, SNB, YR and also APR to YR. In addition, several markers in this chromosomal region (0cM, 8.9cM) were significantly associated with both seedling and APR to YR and also APR to TS (unpublished results). This is interesting because the '2NS' translocation segment from *A. ventricosa* on the distal end of chromosome 2AS has been previously reported to carry resistance to many diseases: strawbreaker foot rot (eyespot) caused by *Pseudocercosporella herpotrichoides* (*Pch1*) (Doussinault et al., 1983), YR (*Yr17*), stem rust caused by *P. graminis* (*Sr38*), LR (*Lr37*) (Bariana and Mcintosh, 1993), cereal cyst caused by *Heterodera avenae* (*Cre5*) (Jahier et al., 2001), root knot caused by *Meloidogyne* spp. (*Rkn3*) (Williamson et al., 2013) and blast caused by *Magnaporthe oryzae* (Cruz et al., 2016). So, we further explored the 2AS chromosomal region and looked at all the genes in the interval from 0 to 7,123,325 bp where the significant markers were located.

There were 228 genes in this region among which seventeen had disease resistance orthologs and NB-ARC, LRR and/or P-loop containing nucleoside triphosphate hydrolase domains in their transcripts. This included seven genes with disease resistance RPP13-like protein orthologs, four genes with disease resistance protein RGA3 orthologs, two genes with disease resistance protein RGA2 orthologs, two genes with disease resistance protein RPM1 orthologs, one gene with disease resistance protein RXW24L ortholog and one gene with disease resistance protein ortholog (Supplementary Table 5). Among the other genes, those with defensin, PAL and ABCG transporter family member orthologs are interesting as they are also known to be involved in disease resistance. Plant defensins are cysteine-rich peptides involved in plant innate immunity that are generally active against a broad spectrum of fungal pathogens and other microbes

(Broekaert et al., 1995; Carvalho and Gomes, 2009). While the PAL gene has been discussed earlier, Tonnessen et al. (2015) reported a rice PAL gene (*OsPAL4*) that was associated with broad spectrum disease resistance. Finally, the ABCG transporter family (also pleiotropic drug resistance family) members are also known to be involved in plant defense (Stukkens et al., 2005; Stein et al., 2006; Krattinger et al., 2009). We hypothesize that either combinations of *R*-genes or genes that confer broad spectrum resistance are responsible for the multiple disease resistance associated with lines carrying this segment. However, not all the genes in this chromosomal segment might be effective as races MBJ/SP and MCJ/SP are virulent to *Lr37* gene and the Ug99 group of races in Kenya are virulent to *Sr38* gene, both of which are linked to the *Yr17* gene.

CONCLUSION

We have identified several markers and potential candidate genes associated with seedling resistance to LR, YR, SNB and TS and also APR to YR. However, these results should be taken with caution for two reasons: (i) Our ability to identify potential candidate genes is limited by the resolution of the GBS markers and the current reference wheat genome assembly (ii) Given the very high LD in wheat, there could be several hundreds of genes in the location of a significant marker and it is not possible to identify the causal gene with just GWAS. Analyzing the significance and the LD of markers around a significant marker will help to delineate the most likely interval for the causal gene. The genes in that interval have to be narrowed down using fine mapping and the final candidates have to be functionally characterized using gene editing, gene silencing or targeting induced local lesions in genomes (TILLING) populations etc. Nevertheless, GWAS is the first step to identify candidate genes at the population level and can provide valuable information on the genetic architecture of the traits.

Our results support previous findings that plant defense mechanisms against pathogens are multifaceted and complex. While ETI mediated by NBS-LRR genes plays an important role, the defense response genes that govern basal resistance and PTI should not be overlooked. However, some of the rust resistance genes that are present in a significant frequency in CIMMYT germplasm were not identified in this study. This might be due to several reasons: (i) the limited phenotypic variability in these advanced breeding lines that were selected for rust resistance (ii) the high frequency of these genes in the lines (iii) the exclusion of several GBS markers (that could be potentially associated), because their positions were not available in the POPSEQ map. We also observed that several genes were associated with resistance in the same genetic position. In this case, the physical map can provide better insight into those genes. The disease resistance gene rich segment on chromosome 2AS is very promising and should be explored further for use in breeding. We conclude that identifying candidate genes linked to significant markers in GWAS is feasible in hexaploid wheat using GBS markers, POPSEQ map and Ensembl plants, thus creating opportunities for accelerating gene cloning and molecular breeding in wheat.

REFERENCES

Abeysekara, N.S., T.L. Friesen, B. Keller, and J.D. Faris. 2009. Identification and characterization of a novel host-toxin interaction in the wheat-*Stagonospora nodorum* pathosystem. Theor Appl Genet 120: 117–126.

Acharya, B.R., S. Raina, S.B. Maqbool, G. Jagadeeswaran, S.L. Mosher, H.M. Appel, et al. 2007. Overexpression of CRK13, an Arabidopsis cysteine-rich receptor-like kinase, results in enhanced resistance to *Pseudomonas syringae*. Plant J 50(3): 488–499.

Adhikari, T.B., E.W. Jackson, S. Gurung, J.M. Hansen, and J.M. Bonman. 2011. Association

Mapping of Quantitative Resistance to *Phaeosphaeria nodorum* in Spring Wheat Landraces from the USDA National Small Grains Collection. Phytopathology 101(11): 1301–1310.

Arseniuk, E., P.C. Czembor, A. Czaplicki, Q. Song, P.B. Cregan, D.L. Hoffman, et al. 2004. QTL controlling partial resistance to *Stagonospora nodorum* leaf blotch in winter wheat cultivar Alba. Euphytica 137(2): 225–231.

Bansal, U.K., M.J. Hayden, M.B. Gill, and H.S. Bariana. 2009. Chromosomal location of an uncharacterised stripe rust resistance gene in wheat. Euphytica 171: 121–127.

Bariana, H.S., and R.A. Mcintosh. 1993. Cytogenetic studies in wheat. XV. Location of rust resistance genes in VPM1 and their genetic linkage with other disease resistance genes in chromosome 2A. Genome 36: 476–482.

Basnet, B.R., R.P. Singh, S.A. Herrera-Foessel, A.M.H. Ibrahim, J. Huerta‐Espino, V. Calvo-Salazar, et al. 2013. Genetic Analysis of Adult Plant Resistance to Yellow Rust and Leaf Rust in Common Spring Wheat Quaiu 3. Plant Dis 97(6): 728–736.

Baulcombe, D. 2004. RNA silencing in plants. Nature 431: 356–363.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc 57(1): 289–300.

Bhathal, J.S., R. Loughman, and J. Speijers. 2003. Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and Septoria nodorum blotch. Eur J Plant Pathol 109: 435–443.

van der Biezen, E.A., and J.D.G. Jones. 1998. The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr Biol 8(7): R226–R227.

Birk, Y., A. Gertler, and S. Khalef. 1963. A pure trypsin inhibitor from soya beans. Biochem J 87:

281–284.

Bittner-Eddy, P.D., and J.L. Beynon. 2001. The Arabidopsis downy mildew resistance gene, RPP13-Nd, functions independently of NDR1 and EDS1 and does not require the accumulation of salicylic acid. Mol Plant-Microbe Interact 14(3): 416–421.

Bittner-Eddy, P.D., I.R. Crute, E.B. Holub, and J.L. Beynon. 2000. RPP13 is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. Plant J 21(2): 177–188.

Bolser, D., D.M. Staines, E. Pritchard, and P. Kersey. 2016. Ensembl Plants: Integrating Tools for Visualizing, Mining, and Analyzing Plant Genomics Data. p. 115–140. *In* Plant Bioinformatics: Methods and Protocols, Methods in Molecular Biology.

Bostwick, D.E., H.W. Ohm, and G. Shaner. 1993. Inheritance of Septoria Glume Blotch Resistance in Wheat. Crop Sci 33: 439–443.

Bouwmeester, K., M. de Sain, R. Weide, A. Gouget, S. Klamer, H. Canut, et al. 2011. The lectin receptor kinase LecRK-I.9 is a novel *Phytophthora* resistance component and a potential host target for a RXLR effector. PLoS Pathog 7(3): e1001327.

Bowman, D.E. 1946. Differentiation of Soy Bean Antitryptic Factors. Exp Biol Med 63(3): 547–550.

Box, G.E.P., and D.R. Cox. 1964. An Analysis of Transformations. J R Stat Soc Ser B 26(2): 211–252.

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23(19): 2633–2635.

Breseghello, F., and M.E. Sorrells. 2006. Association mapping of kernel size and milling quality

in wheat *(Triticum aestivum* L.) cultivars. Genetics 172(2): 1165–1177.

Broekaert, W.F., F.R. Terras, B.P. Cammue, and R.W. Osborn. 1995. Plant defensins: novel antimicrobial peptides as components of the host defense system. Plant Physiol 108: 1353–1358.

Brown-Guedira, G.L., S. Singh, and A.K. Fritz. 2003. Performance and Mapping of Leaf Rust Resistance Transferred to Wheat from Triticum timopheevii subsp. armeniacum. Phytopathology 93(7): 784–789.

Brutus, A., F. Sicilia, A. Macone, F. Cervone, G. De Lorenzo, and G. De Lorenzo. 2010. A domain swap approach reveals a role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. Proc Natl Acad Sci U S A 107(20): 9452–7.

Bryan, J., A. Munoz, X. Zhang, M. Dufer, G. Drews, P. Krippeit-Drews, et al. 2007. ABCC8 and ABCC9: ABC transporters that regulate K+ channels. Pflugers Arch Eur J Physiol 453(5): 703–718.

La Camera, S., G. Gouzerh, S. Dhondt, L. Hoffmann, B. Fritig, M. Legrand, et al. 2004. Metabolic reprogramming in plant innate immunity: The contributions of phenylpropanoid and oxylipin pathways. Immunol Rev 198: 267–284.

Carvalho, A. de O., and V.M. Gomes. 2009. Plant defensins-Prospects for the biological functions and biotechnological properties. Peptides 30(5): 1007–1020.

Césari, S., H. Kanzaki, T. Fujiwara, M. Bernoux, V. Chalvon, Y. Kawano, et al. 2014. The NB-LRR proteins RGA4 and RGA5 interact functionally and physically to confer disease resistance. EMBO J 33(17): 1941–1959.

Chapman, J.A., M. Mascher, A. Buluç, K. Barry, E. Georganas, A. Session, et al. 2015. A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome.

Genome Biol 16(1): 26.

Chen, Z. 2001. A Superfamily of Proteins with Novel Cysteine-Rich Repeats. Plant Physiol 126(2): 473–476.

Chen, X.M. 2005. Epidemiology and control of stripe rust [Puccinia striiformis f . sp . tritici] on wheat. Can J Plant Pathol 27: 314–337.

Chhetri, M., H. Bariana, P. Kandiah, and U. Bansal. 2016. *Yr58* : A New Stripe Rust Resistance Gene and Its Interaction with *Yr46* for Enhanced Resistance. Phytopathology: PHYTO-04-16-018.

Cho, M., Z.-W. Lee, and H.-T. Cho. 2012. ATP-binding cassette B4, an auxin-efflux transporter, stably associates with the plasma membrane and shows distinctive intracellular trafficking from that of PIN-FORMED proteins. Plant Physiol 159(2): 642–54.

Choi, J.J., S.J. Klosterman, and L.A. Hadwiger. 2004. A Promoter from Pea Gene DRR206 Is Suitable to Regulate an Elicitor-Coding Gene and Develop Disease Resistance. Phytopathology 94(6): 651–660.

Corbin, D.R., N. Sauer, and C.J. Lamb. 1987. Differential Regulation of of a Hydroxyproline-Rich Glycoprotein Gene Family in Wounded and Infected Plants. Mol Cell Biol 7(12): 4337–4344.

Cox, T.S., W.J. Raupp, and B.S. Gill. 1994a. Leaf Rust-Resistance Genes *Lr41*, *Lr42*, and *Lr43* Transferred from *Triticum tauschii* to Common Wheat. Crop Sci 34(2): 339–343.

Cox, T.S., R.G. Sears, B.S. Gill, and E.N. Jellen. 1994b. Registration of KS91WGRC11, KS92WGRC15, and KS92WGRC23 Leaf Rust-Resistant Hard Red Winter Wheat Germplasms. Crop Sci 34: 546–547.

Craig, A., R. Ewan, J. Mesmar, V. Gudipati, and A. Sadanandom. 2009. E3 ubiquitin ligases and plant innate immunity. J Exp Bot 60(4): 1123–1132.

Crook, A.D., T.L. Friesen, Z.H. Liu, P.S. Ojiambo, and C. Cowger. 2012. Novel Necrotrophic Effectors from *Stagonospora nodorum* and Corresponding Host Sensitivities in Winter Wheat Germplasm in the Southeastern United States. Phytopathology 102: 498–505.

Crossa, J., J. Burgueño, S. Dreisigacker, M. Vargas, S.A. Herrera-Foessel, M. Lillemo, et al. 2007. Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics 177(3): 1889–1913.

Cruz, C.D., G.L. Peterson, W.W. Bockus, P. Kankanala, J. Dubcovsky, K.W. Jordan, et al. 2016. The 2NS Translocation from *Aegilops ventricosa* Confers Resistance to the *Triticum* Pathotype of *Magnaporthe oryzae*. Crop Sci 0(0): 0.

Dangl, J.L., and J.D. Jones. 2001. Plant pathogens and integrated defence responses to infection. Nature 411(6839): 826–833.

Daudi, A., Z. Cheng, J.A.O. Brien, N. Mammarella, S. Khan, F.M. Ausubel, et al. 2012. The Apoplastic Oxidative Burst Peroxidase in *Arabidopsis* Is a Major Component of Pattern-Triggered Immunity. Plant Cell 24(1): 275–287.

Davenport, R. 2002. Glutamate receptors in plants. Ann Bot 90(5): 549–557.

Debener, T., H. Lehnackers, M. Arnold, and J.L. Dangl. 1991. Identification and molecular mapping of a single *Arabidopsis thaliana* locus determining resistance to a phytopathogenic *Pseudomonas syringae* isolate. Plant J 1(3): 289–302.

Decreux, A., and J. Messiaen. 2005. Wall-associated kinase WAK1 interacts with cell wall pectins in a calcium-induced conformation. Plant Cell Physiol 46(2): 268–278.

Delannoy, E., W.A. Stanley, C.S. Bond, and I.D. Small. 2007. Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. Biochem Soc Trans 35(Pt 6): 1643–1647.

Dielen, A.S., S. Badaoui, T. Candresse, and S. German-Retana. 2010. The ubiquitin/26S proteasome system in plant-pathogen interactions: A never-ending hide-and-seek game. Mol Plant Pathol 11(2): 293–308.

Dixon, R.A., L. Achnine, P. Kota, C.J. Liu, M.S.S. Reddy, and L. Wang. 2002. The phenylpropanoid pathway and plant defence - A genomics perspective. Mol Plant Pathol 3(5): 371–390.

Doussinault, G., A. Delibes, R. Sanchez-Monge, and F. Garcia-Olmedo. 1983. Transfer of a dominant gene for resistance to eyespot disease from a wild grass to hexaploid wheat. Nature 303(5919): 698–700.

Du, C.G., L.R. Nelson, and M.E. McDaniel. 1999. Diallel analysis of gene effects conditioning resistance to *Stagonospora nodorum* (Berk.) in wheat. Crop Sci 39(3): 686–690.

Dyck, P.L., and E.R. Kerber. 1977. Inheritance of leaf rust resistance in wheat cultivars Rafaela and EAP 26127 and chromosome location of gene *Lr17*. Can J Genet Cytol 19: 355–358.

Dyck, P.L., and E.R. Kerber. 1981. Aneuploid analysis of a gene for leaf rust resistance derived from the common wheat cultivar Terenzio. Can J Genet Cytol 23(969): 405–409.

Ederli, L., L. Madeo, O. Calderini, C. Gehring, C. Moretti, R. Buonaurio, et al. 2011. The *Arabidopsis thaliana* cysteine-rich receptor-like kinase CRK20 modulates host responses to *Pseudomonas syringae* pv. *tomato* DC3000 infection. J Plant Physiol 168(15): 1784–1794.

Edwards, K., C.L. Cramer, G.P. Bolwell, R.A. Dixon, W. Schuch, and C.J. Lamb. 1985. Rapid transient induction of phenylalanine ammonia-lyase mRNA in elicitor-treated bean cells. Proc Natl Acad Sci U S A 82(20): 6731–6735.

Effertz, R.J., S.W. Meinhardt, J.A. Anderson, J.G. Jordahl, and L.J. Francl. 2002. Identification of a Chlorosis-Inducing Toxin from *Pyrenophora tritici-repentis* and the Chromosomal

Location of an Insensitivity Locus in Wheat. Phytopathology 92(5): 527–533.

Ellis, J.G., E.S. Lagudah, W. Spielmeyer, and P.N. Dodds. 2014. The past, present and future of breeding rust resistant wheat. Front Plant Sci 5(November): 641.

Endelman, J.B. 2011. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. Plant Genome 4: 250–255.

Endelman, J.B., and J.-L. Jannink. 2012. Shrinkage estimation of the realized relationship matrix. G3 Genes|Genomes|Genetics 2(11): 1405–1413.

Eyal, Z., A.L. Scharen, J.M. Prescott, and M. van Ginkel. 1987. The Septoria Diseases of Wheat: Concepts and methods of disease management. CIMMYT., Mexico, D.F.

Faris, J.D., J.A. Anderson, L.J. Francl, and J.G. Jordahl. 1996. Chromosomal location of a gene conditioning insensitivity in wheat to a necrosis-inducing culture filtrate from *Pyrenophora tritici-repentis*. Phytopathology 86(5): 459–463.

Faris, J.D., and T.L. Friesen. 2005. Identification of quantitative trait loci for race-nonspecific resistance to tan spot in wheat. Theor Appl Genet 111(2): 386–392.

Faris, J.D., Z. Zhang, H. Lu, S. Lu, L. Reddy, S. Cloutier, et al. 2010. A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. Proc Natl Acad Sci U S A 107(30): 13544–13549.

Faris, J.D., Z. Zhang, J.B. Rasmussen, and T.L. Friesen. 2011. Variable Expression of the *Stagonospora nodorum* Effector SnToxA Among Isolates Is Correlated with Levels of Disease in Wheat. Mol Plant-Microbe Interact 24(12): 1419–1426.

Feng, J., H. Ma, and G.R. Hughes. 2004. Genetics of resistance to Stagonospora nodorum blotch of hexaploid wheat. Crop Sci 44: 2043–2048.

Feuillet, C., S. Travella, N. Stein, L. Albar, A. Nublat, and B. Keller. 2003. Map-based isolation

of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (*Triticum aestivum* L.) genome. Proc Natl Acad Sci U S A 100(25): 15253–8.

Flor, H.H. 1956. The complementary genic systems in flax and flax rust. Adv Genet 8: 29–54.

Footitt, S., S.P. Slocombe, V. Larner, S. Kurup, Y. Wu, T. Larson, et al. 2002. Control of germination and lipid mobilization by COMATOSE, the Arabidopsis homologue of human ALDP. EMBO J 21(12): 2912–2922.

Fried, P.M., and E. Meister. 1987. Inheritance of Leaf and Head Resistance of Winter Wheat to *Septoria nodorum* in a Diallel Cross. Phytopathology 77(10): 1371–1375.

Friesen, T.L., C. Chu, S.S. Xu, and J.D. Faris. 2012. SnTox5-*Snn5*: A novel *Stagonospora nodorum* effector-wheat gene interaction and its relationship with the SnToxA-*Tsn1* and SnTox3-*Snn3-B1* interactions. Mol Plant Pathol 13: 1101–1109.

Friesen, T.L., and J.D. Faris. 2004. Molecular mapping of resistance to *Pyrenophora tritici-repentis* race 5 and sensitivity to Ptr ToxB in wheat. TAG Theor Appl Genet 109(3): 464–471.

Friesen, T.L., J.D. Faris, P.S. Solomon, and R.P. Oliver. 2008a. Host-specific toxins: Effectors of necrotrophic pathogenicity. Cell Microbiol 10(7): 1421–1428.

Friesen, T.L., S.W. Meinhardt, and J.D. Faris. 2007. The *Stagonospora nodorum*-wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. Plant J 51: 681–692.

Friesen, T.L., E.H. Stukenbrock, Z. Liu, S. Meinhardt, H. Ling, J.D. Faris, et al. 2006. Emergence of a new disease as a result of interspecific virulence gene transfer. Nat Genet 38(8): 953–956.

Friesen, T.L., Z. Zhang, P.S. Solomon, R.P. Oliver, and J.D. Faris. 2008b. Characterization of the

interaction of a novel *Stagonospora nodorum* host-selective toxin with a wheat susceptibility gene. Plant Physiol 146: 682–693.

Gao, Y., J.D. Faris, Z. Liu, Y.M. Kim, R.A. Syme, R.P. Oliver, et al. 2015. Identification and Characterization of the SnTox6-Snn6 Interaction in the Parastagonospora nodorum – Wheat Pathosystem. Mol Plant-Microbe Interact 28(5): 615–625.

García-Olmedo, F., A. Molina, A. Segura, and M. Moreno. 1995. The defensive role of nonspecific lipid-transfer proteins in plants. Trends Microbiol 3(2): 72–74.

Geisler, M., J.J. Blakeslee, R. Bouchard, O.R. Lee, V. Vincenzetti, A. Bandyopadhyay, et al. 2005. Cellular efflux of auxin catalyzed by the Arabidopsis MDR/PGP transporter AtPGP1. Plant J 44(2): 179–194.

Gómez-Gómez, L., and T. Boller. 2000. FLS2: An LRR Receptor–like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in *Arabidopsis*. Mol Cell 5(6): 1003–1011.

Greenberg, J.T., and N. Yao. 2004. The role of regulation of programmed cell death in plant-pathogen interactions. Cell Microbiol 6(3): 201–211.

Guo, Y.H., Y.P. Yu, D. Wang, C.A. Wu, G.D. Yang, J.G. Huang, et al. 2009. GhZFP1, a novel CCCH-type zinc finger protein from cotton, enhances salt stress tolerance and fungal disease resistance in transgenic tobacco by interacting with GZIRD21A and GZIPR5. New Phytol 183(1): 62–75.

Gupta, O.P., V. Permar, V. Koundal, U.D. Singh, and S. Praveen. 2012. MicroRNA regulated defense responses in *Triticum aestivum* L. during *Puccinia graminis* f.sp. *tritici* infection. Mol Biol Rep 39(2): 817–824.

Hammond-Kosack, K.E., and J.D.G. Jones. 1997. Plant disease resistance genes. Annu Rev Plant Physiol Plant Mol Biol 48: 575–607.

He, Z.H., I. Cheeseman, D. He, and B.D. Kohorn. 1999. A cluster of five cell wall-associated receptor kinase genes, *Wak1-5*, are expressed in specific organs of *Arabidopsis*. Plant Mol Biol 39(6): 1189–1196.

He, Z., M. Fujiki, and B.D. Kohorn. 1996. A Cell Wall-associated, Receptor-like Protein Kinase. J Biol Chem 271(33): 19789–19793.

He, Z.H., D. He, and B.D. Kohorn. 1998. Requirement for the induced expression of a cell wall-associated receptor kinase for survival during the pathogen response. Plant J 14(1): 55–63.

Hiebert, C.W., J.B. Thomas, B.D. McCallum, and D.J. Somers. 2008. Genetic mapping of the wheat leaf rust resistance gene *Lr60* (*LrW2*). Crop Sci 48(3): 1020–1026.

Higgins, C.F. 1992. ABC Transporters: From Microorganisms to Man. Annu Rev Cell Biol 8: 67–113.

Huang, L., S. a Brooks, W. Li, J.P. Fellers, H.N. Trick, and B.S. Gill. 2003. Map-Based Cloning of Leaf Rust Resistance Gene *Lr21* From the Large and Polyploid Genome of Bread Wheat. Genetics 164: 655–664.

Huerta-Espino, J., R.P. Singh, S. Germán, B.D. McCallum, R.F. Park, W.Q. Chen, et al. 2011. Global status of wheat leaf rust caused by *Puccinia triticina*. Euphytica 179(1): 143–160.

International Wheat Genome Sequencing Consortium (IWGSC). 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345(6194): 1251788.

Jahier, J., P. Abelard, A.M. Tanguy, F. Dedryver, R. Rivoal, S. Khatkar, et al. 2001. The Aegilops ventricosa segment on chromosome 2AS of the wheat cultivar "VPM1" carries the cereal cyst nematode resistance gene Cre5. Plant Breed 120(2): 125–128.

Johnson, R. 1984. A critical analysis of durable resistance. Annu Rev Phytopathol 22: 309–330.

Jones, J.D.G., and J.L. Dangl. 2006. The plant immune system. Nature 444: 323–329.

Jordá, L., and P. Vera. 2000. Local and systemic induction of two defense-related subtilisin-like protease promoters in transgenic Arabidopsis plants. Luciferin induction of PR gene expression. Plant Physiol 124(3): 1049–1058.

Jørgensen, J.H. 1992. Discovery, characterization and exploitation of Mlo powdery mildew resistance in barley. Euphytica 63: 141–152.

Juliana, P., J.E. Rutkoski, J.A. Poland, R.P. Singh, S. Murugasamy, S. Natesan, et al. 2015. Genome-Wide Association Mapping for Leaf Tip Necrosis and Pseudo-black Chaff in Relation to Durable Rust Resistance in Wheat. Plant Genome 8(2).

Kader, J. 1996. Lipid-Transfer Proteins in Plants. Annu Rev Plant Physiol Plant Mol Biol 47: 627–654.

Kang, S., H.B. Kim, H. Lee, J.Y. Choi, S. Heu, C.J. Oh, et al. 2006. Overexpression in *Arabidopsis* of a plasma membrane-targeting glutamate receptor from small radish increases glutamate-mediated Ca2+ influx and delays fungal infection. Mol Cells 21(3): 418–427.

Kawano, T. 2003. Roles of the reactive oxygen species-generating peroxidase reactions in plant defense and growth induction. Plant Cell Rep 21(9): 829–837.

Kazan, K., and J.M. Manners. 2009. Linking development to defense: auxin in plant-pathogen interactions. Trends Plant Sci 14(7): 373–382.

Kim, S.A., J.M. Kwak, S.K. Jae, M.H. Wang, and H.G. Nam. 2001. Overexpression of the *AtGluR2* gene encoding an arabidopsis homolog of mammalian glutamate receptors impairs calcium utilization and sensitivity to ionic stress in transgenic plants. Plant Cell Physiol 42(1): 74–84.

King, J.E., R.J. Cook, and S.C. Melville. 1983. A review of *Septoria* diseases of wheat and barley.

Ann Appl Biol 103: 345–373.

Kolmer, J.A., X. Chen, and Y. Jin. 2009. Diseases which challenge global wheat production - the wheat rusts. *In* Carver, B.F. (ed.), Wheat Science and Trade. Wiley-Blackwell, Ames, IA.

Krattinger, S.G., E.S. Lagudah, W. Spielmeyer, R.P. Singh, J. Huerta-espino, H. Mcfadden, et al. 2009. A Putative ABC Transporter Confers Durable Resistance to Multiple Fungal Pathogens in Wheat. Science 323: 1360–1363.

Kristensen, A.K., J. Brunstedt, K.K. Nielsen, P. Roepstorff, and J.D. Mikkelsen. 2000. Characterization of a new antifungal non-specific lipid transfer protein (nsLTP) from sugar beet leaves. Plant Sci 155(1): 31–40.

Kuhar, K., R. Kansal, B. Subrahmanyam, K.R. Koundal, K. Miglani, and V.K. Gupta. 2013. A Bowman-Birk protease inhibitor with antifeedant and antifungal activity from *Dolichos biflorus*. Acta Physiol Plant 35(6): 1887–1903.

Kuraparthy, V., P. Chhuneja, H.S. Dhaliwal, S. Kaur, R.L. Bowden, and B.S. Gill. 2007. Characterization and mapping of cryptic alien introgression from *Aegilops geniculata* with new leaf rust and stripe rust resistance genes *Lr57* and *Yr40* in wheat. Theor Appl Genet 114: 1379–1389.

Lagudah, E.S., S.G. Krattinger, S. Herrera-Foessel, R.P. Singh, J. Huerta-Espino, W. Spielmeyer, et al. 2009. Gene-specific markers for the wheat gene *Lr34/Yr18/Pm38* which confers resistance to multiple fungal pathogens. Theor Appl Genet 119(5): 889–898.

Laluk, K., S. AbuQamar, and T. Mengiste. 2011. The Arabidopsis Mitochondria-Localized Pentatricopeptide Repeat Protein PGN Functions in Defense against Necrotrophic Fungi and Abiotic stress tolerance. Plant Physiol 156(4): 2053–2068.

Lam, H.-M., J. Chiu, M.-H. Hsieh, L. Meisel, I.C. Oliveira, M. Shin, et al. 1998. Glutamate-

receptor genes in plants. Nature 396: 125–126.

Lamari, L., and C.C. Bernier. 1989. Evaluation of wheat lines and cultivars to tan spot [*Pyrenophora tritici-repentis*] based on lesion type. Can J Plant Pathol 11: 49–56.

Lamport, D.T.A., and D.H. Northcote. 1960. Hydroxyproline in primary cell walls of higher plants. Nature 188: 665–666.

Li, Y., Y.-G. Lu, Y. Shi, L. Wu, Y.-J. Xu, F. Huang, et al. 2014. Multiple rice microRNAs are involved in immunity against the blast fungus *Magnaporthe oryzae*. Plant Physiol 164: 1077–1092.

Li, F., J. Wang, C. Ma, Y. Zhao, Y. Wang, A. Hasi, et al. 2013. Glutamate Receptor-Like Channel3.3 Is Involved in Mediating Glutathione-Triggered Cytosolic Calcium Transients, Transcriptional Changes, and Innate Immunity Responses in Arabidopsis. Plant Physiol 162: 1497–1509.

Liu, Z.H., J.D. Faris, S.W. Meinhardt, S. Ali, J.B. Rasmussen, and T.L. Friesen. 2004a. Genetic and Physical Mapping of a Gene Conditioning Sensitivity in Wheat to a Partially Purified Host-Selective Toxin Produced by *Stagonospora nodorum*. Phytopathology 94(5): 1056–1060.

Liu, Z., J.D. Faris, R.P. Oliver, K.C. Tan, P.S. Solomon, M.C. McDonald, et al. 2009. SnTox3 acts in effector triggered susceptibility to induce disease on wheat carrying the *Snn3* gene. PLoS Pathog 5(9): e1000581.

Liu, Z., T.L. Friesen, H. Ling, S.W. Meinhardt, R.P. Oliver, J.B. Rasmussen, et al. 2006. The *Tsn1*-ToxA interaction in the wheat-*Stagonospora nodorum* pathosystem parallels that of the wheat-tan spot system. Genome 49: 1265–1273.

Liu, Z.H., T.L. Friesen, J.B. Rasmussen, S. Ali, S.W. Meinhardt, and J.D. Faris. 2004b.

Quantitative Trait Loci Analysis and Mapping of Seedling Resistance to Stagonospora nodorum Leaf Blotch in Wheat. Phytopathology 94(10): 1061–1067.

Liu, Z., Z. Zhang, J.D. Faris, R.P. Oliver, R. Syme, M.C. McDonald, et al. 2012. The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring *Snn1*. PLoS Pathog 8(1): e1002467.

Llorente, F., P. Muskett, A. Sánchez-Vallet, G. López, B. Ramos, C. Sánchez-Rodríguez, et al. 2008. Repression of the auxin response pathway increases *Arabidopsis* susceptibility to necrotrophic fungi. Mol Plant 1(3): 496–509.

van Loon, L.C., and E. a van Strien. 1999. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol Mol Plant Pathol 55: 85–97.

Mackey, D., B.F. Holt, A. Wiig, and J.L. Dangl. 2002. RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. Cell 108: 743–754.

Maldonado, A.M., P. Doerner, R.A. Dixon, C.J. Lamb, and R.K. Cameron. 2002. A putative lipid transfer protein involved in systemic resistance signalling in *Arabidopsis*. Nature 419: 399–403.

Mammarella, N.D., Z. Cheng, Z.Q. Fu, A. Daudi, G.P. Bolwell, X. Dong, et al. 2015. Apoplastic peroxidases are required for salicylic acid-mediated defense against *Pseudomonas syringae*. Phytochemistry 112: 110–121.

Marais, F., A. Marais, B. McCallum, and Z. Pretorius. 2009. Transfer of leaf rust and stripe rust resistance genes *Lr62* and *Yr42* from *Aegilops neglecta* req. ex bertol. to common wheat. Crop Sci 49: 871–879.

Marais, G.F., B. McCallum, and A.S. Marais. 2006. Leaf rust and stripe rust resistance genes derived from *Aegilops sharonensis*. Euphytica 149: 373–380.

Marais, G.F., B. Mccallum, J. Snyman, Z.A. Pretorius, and A.S. Marais. 2005. Leaf rust and stripe rust resistance genes *Lr54* and *Yr37* transferred to wheat from *Aegilops kotschyi*. Plant Breed 124: 538–541.

Marasas, C.N., M. Smale, and R.P. Singh. 2004. The Economic Impact in Developing Countries of Leaf Rust Resistance Breeding in CIMMYT-Related spring bread wheat. CIMMYT Mex DF.

McDonald, B. a, and C. Linde. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annu Rev Phytopathol 40: 349–379.

Mcintosh, R.A., J. Dubcovsky, W.J. Rogers, C.F. Morris, R. Appels, and X.C. Xia. 2009. Catalogue of gene symbols for wheat: 2009 supplement.

Mcintosh, R.A., J. Dubcovsky, W.J. Rogers, C.F. Morris, R. Appels, and X.C. Xia. 2014. Catalogue of gene symbols for wheat: 2013-2014 supplement.

McIntosh, R.A., J. Dubcovsky, W. Rogers, C. Morris, R. Appels, and X.C. Xia. 2016. Catalogue of gene symbols for wheat: 2015-2016 supplement.

McIntosh, R.A., B. Friebe, J. Jiang, D. The, and B.S. Gill. 1995. Cytogenetical studies in wheat XVI. Chromosome location of a new gene for resistance to leaf rust in a Japanese wheat-rye translocation line. Euphytica 82: 141–147.

McIntosh, R.A., T.E. Miller, and V. Chapman. 1982. Cytogenetical studies in wheat XII. *Lr28* for resistance to *Puccinia recondita* and *Sr34* for resistance to *P. graminis tritici*. Z Pflanzenzuchtg 89: 295–306.

McNeal, F.H., C.F. Konzak, E.P. Smith, W.S. Tate, and T.S. Russell. 1971. A uniform system for

recording and processing cereal research data. Agricultural Research Service, United States Department of Agriculture, [Beltsville Md.].

Miao, M., X. Niu, J. Kud, X. Du, J. Avila, T.P. Devarenne, et al. 2016. The ubiquitin ligase SEVEN IN ABSENTIA (SINA) ubiquitinates a defense-related NAC transcription factor and is involved in defense signaling. New Phytol 211: 138–148.

Mohler, V., D. Singh, C. Singrun, and R.F. Park. 2012. Characterization and mapping of *Lr65* in spelt wheat "Altgold Rotkorn." Plant Breed 131: 252–257.

Molina, A., A. Segura, and F. Garcia-Olmedo. 1993. Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant pathogens. FEBS Lett 316(2): 119–122.

Morant, A.V., K. Jørgensen, C. Jørgensen, S.M. Paquette, R. Sánchez-Pérez, B.L. Møller, et al. 2008. β-Glucosidases as detonators of plant chemical defense. Phytochemistry 69(9): 1795–1813.

Morris, E.R., and J.C. Walker. 2003. Receptor-like protein kinases: The keys to response. Curr Opin Plant Biol 6(4): 339–342.

Mudgil, Y., S.-H. Shiu, S.L. Stone, J.N. Salt, and D.R. Goring. 2004. A large complement of the predicted Arabidopsis ARM repeat proteins are members of the U-box E3 ubiquitin ligase family. Plant Physiol 134(1): 59–66.

Murray, G.M., and J.P. Brennan. 2009. Estimating disease losses to the Australian wheat industry. Australas Plant Pathol 38(6): 558–570.

Nisius, A. 1988. The stromacentre in Avena plastids: an aggregation of *β*-glucosidase responsible for the activation of oat-leaf saponins. Planta 173: 474–481.

Noh, B., A.S. Murphy, and E.P. Spalding. 2001. Multidrug Resistance – like Genes of Arabidopsis

Required for Auxin Transport and Auxin-Mediated Development. Plant Cell 13: 2441–2454.

Nusslein-Volhard, C., and E. Wieschaus. 1980. Mutations affecting segment number and polarity in Drosophila. Nature 287: 795–801.

Park, Y.J., H.J. Lee, K.J. Kwak, K. Lee, S.W. Hong, and H. Kang. 2014. MicroRNA400-guided cleavage of pentatricopeptide repeat protein mRNAs renders *Arabidopsis thaliana* more susceptible to pathogenic bacteria and fungi. Plant Cell Physiol 55(9): 1660–1668.

Pearce, G., Y. Yamaguchi, G. Barona, and C.A. Ryan. 2010. A subtilisin-like protein from soybean contains an embedded, cryptic signal that activates defense-related genes. Proc Natl Acad Sci U S A 107(33): 14921–14925.

Pellegrini, L., O. Rohfritsch, B. Fritig, and M. Legrand. 1994. Phenylalanine Ammonia-Lyase in Tobacco. Plant Physiol 106: 877–886.

Peterson, R.F., A.B. Campbell, and A.E. Hannah. 1948. A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. Can J Res 26c(5): 496–500.

Poland, J.A., P.J. Brown, M.E. Sorrells, and J.L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7(2): e32253.

Price, A.L., N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38(8): 904–909.

Qu, L., J. Chen, M. Liu, N. Pan, H. Okamoto, Z. Lin, et al. 2003. Molecular Cloning and Functional Analysis of a Novel Type of Bowman-Birk Inhibitor Gene Family in Rice. Plant Physiol 133: 560–570.

Ralph, S., J.Y. Park, J. Bohlmann, and S.D. Mansfield. 2006. Dirigent proteins in conifer defense:

gene discovery, phylogeny, and differential wound- and insect-induced expression of a family of DIR and DIR-like genes in spruce (*Picea* spp.). Plant Mol Biol 60: 21–40.

Randhawa, M.S., H.S. Bariana, R. Mago, and U.K. Bansal. 2015. Mapping of a new stripe rust resistance locus *Yr57* on chromosome 3BS of wheat. Mol Breed 35(2).

Rawal, H.C., N.K. Singh, and T.R. Sharma. 2013. Conservation, divergence, and genome-wide distribution of PAL and POX A gene families in plants. Int J Genomics 678969.

Rea, P.A. 2007. Plant ATP-Binding Cassette Transporters. Annu Rev Plant Biol 58: 347–375.

Rees, R.R., G. Platz, and R. Mayer. 1982. Yield losses in wheat from yellow spot: comparison of estimates derived from single tillers and plots. Aust J Agric Res 33: 899–908.

Regente, M.C., a M. Giudici, J. Villalaín, and L. de la Canal. 2005. The cytotoxic properties of a plant lipid transfer protein involve membrane permeabilization of target cells. Lett Appl Microbiol 40: 183–189.

Riggleman, B., E. Wieschaus, and P. Schedl. 1989. Molecular analysis of the armadillo locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a Drosophila segment polarity gene. Genes Dev 3: 96–113.

Roelfs, A.P., R.P. Singh, and E.E. Saari. 1992. Rust Diseases of Wheat: Concepts and methods of disease management. CIMMYT, Mexico, D.F.

Rowland, G.G., and E.R. Kerber. 1974. Telocentric mapping in hexaploid wheat of genes for Leaf Rust Resistance and Other Characters Derived From *Aegilops Squarrosa*. Can J Genet Cytol 16: 137–144.

Sarowar, S., Y.J. Kim, K.D. Kim, B.K. Hwang, S.H. Ok, and J.S. Shin. 2009. Overexpression of lipid transfer protein (LTP) genes enhances resistance to plant pathogens and LTP functions in long-distance systemic signaling in tobacco. Plant Cell Rep 28(3): 419–427.

Seevers, P.M., and J.M. Daly. 1970. Studies on wheat stem rust resistance controlled at the *Sr6* locus. II. Peroxidase activities. Phytopathology 60: 1642–1647.

Segura, A., M. Moreno, and F. Garcia-Olmedo. 1993. Purification and antipathogenic activity of lipid transfer proteins (LTPs) from the leaves of *Arabidopsis* and spinach. FEBS Lett 332(3): 243–246.

Shabeer, A., and W.W. Bockus. 1988. Tan spot effects on yield and yield components relative to growth stage in winter wheat. Plant Dis 72: 599–602.

Shadle, G.L., S.V. Wesley, K.L. Korth, F. Chen, C. Lamb, and R.A. Dixon. 2003. Phenylpropanoid compounds and disease resistance in transgenic tobacco with altered expression of L-phenylalanine ammonia-lyase. Phytochemistry 64(1): 153–161.

Shailasree, S., K.R. Kini, S. Deepak, B.S. Kumudini, and H.S. Shetty. 2004. Accumulation of hydroxyproline-rich glycoproteins in pearl millet seedlings in response to Sclerospora graminicola infection. Plant Sci 167: 1227–1234.

Shi, G., T.L. Friesen, J. Saini, S.S. Xu, J.B. Rasmussen, and J.D. Faris. 2015. The Wheat Gene *Snn7* Confers Susceptibility on Recognition of the *Parastagonospora nodorum* Necrotrophic Effector SnTox7. Plant Genome 8(2).

Shipton, W.A., W.R.J. Boyd, A.A. Rosielle, and B.I. Shearer. 1971. The common Septoria diseases of wheat. Bot Rev 37(2): 231–262.

Shiu, S., and A.B. Bleecker. 2001. Plant Receptor-Like Kinase Gene Family : Diversity, Function, and Signaling. Sci stke 113(re22): 1–13.

Showalter, A.M., J.N. Bell, C.L. Cramer, J.A. Bailey, J.E. Varner, and C.J. Lamb. 1985. Accumulation of hydroxyproline-rich glycoprotein mRNAs in response to fungal elicitor and infection. Proc Natl Acad Sci 82: 6551–6555.

Siezen, R.J., and J.A. Leunissen. 1997. Subtilases: the superfamily of subtilisin-like serine proteases. Protein Sci 6: 501–523.

Singh, P.K., E. Duveiller, and R.P. Singh. 2011. Evaluation of CIMMYT germplasm for resistance to leaf spotting diseases of wheat. Czech J Genet Plant Breed 47: S102–S108.

Singh, P.K., J.M. Gonzalez-Hernandez, and S.B. Goodwin. 2006a. Identification and Molecular Mapping of a Gene Conferring Resistance to *Pyrenophora tritici-repentis* race 3 in tetraploid wheat. Phytopathology 96(8): 885–889.

Singh, R.P., J. Huerta-Espino, and H.M. William. 2005. Genetics and breeding for durable resistance to leaf and stripe rusts in wheat. Turkish J Agric For 29: 121–127.

Singh, P., Y.-C. Kuo, S. Mishra, C.-H. Tsai, C.-C. Chien, C.-W. Chen, et al. 2012. The Lectin Receptor Kinase-VI.2 Is Required for Priming and Positively Regulates *Arabidopsis* Pattern-Triggered Immunity. Plant Cell 24(3): 1256–1270.

Singh, R.P., and R.A. McIntosh. 1984. Complementary genes for resistance to *Puccinia recondita tritici* in *Triticum aestivum* II. Cytogenetic studies. Can J Genet Cytol 26: 736–742.

Singh, P.K., M. Mergoum, T.B. Adhikari, S.F. Kianian, and E.M. Elias. 2006b. Chromosomal location of genes for seedling resistance to tan spot and Stagonospora nodorum blotch in tetraploid wheat. Euphytica 155: 27–34.

Singh, P.K., M. Mergoum, J.L. Gonzalez-Hernandez, S. Ali, T.B. Adhikari, S.F. Kianian, et al. 2008. Genetics and molecular mapping of resistance to necrosis inducing race 5 of *Pyrenophora tritici-repentis* in tetraploid wheat. Mol Breed 21: 293–304.

Singh, P.K., R.P. Singh, E. Duveiller, M. Mergoum, T.B. Adhikari, and E.M. Elias. 2009. Genetics of wheat-*Pyrenophora tritici*-repentis interactions. Euphytica 171(1): 1–13.

Small, I.D., and N. Peeters. 2000. The PPR motif - a TPR-related motif prevalent in plant

organellar proteins. Trends Biochem Sci 25(2): 46–47.

Somers, D.J., P. Isaac, and K. Edwards. 2004. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theor Appl Genet 109: 1105–1114.

Song, W.Y., G.L. Wang, L.L. Chen, H.S. Kim, L.Y. Pi, T. Holsten, et al. 1995. A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270: 1804–1806.

Southerton, S., and B. Deverall. 1990. Changes in phenylalanine ammonia-lyase and peroxidase activities in wheat cultivars expressing resistance to the leaf-rust fungus. Plant Pathol 39: 223–230.

Stein, M., J. Dittgen, C. Sanchez-Rodriguez, B. Hou, A. Molina, P. Schulze-lefert, et al. 2006. Arabidopsis PEN3/PDR8, an ATP Binding Cassette Transporter, Contributes to Nonhost Resistance to Inappropriate Pathogens That Enter by Direct Penetration. Plant Cell 18: 731–746.

Stukkens, Y., A. Bultreys, T. Trombik, D. Vanham, and M. Boutry. 2005. NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from *Nicotiana plumbaginifolia*, plays a major role in plant pathogen defense. Plant Physiol 139: 341–352.

Sun, X., G. Bai, B.F. Carver, and R. Bowden. 2010. Molecular Mapping of Wheat Leaf Rust Resistance Gene Lr42. Crop Sci 50: 59–66.

Suzuki, H., S. Takahashi, R. Watanabe, Y. Fukushima, N. Fujita, A. Noguchi, et al. 2006. An isoflavone conjugate-hydrolyzing β-glucosidase from the roots of soybean (*Glycine max*) seedlings. J Biol Chem 281(40): 30251–30259.

Tadesse, W., S.L.K. Hsam, G. Wenzel, and F.J. Zeller. 2006a. Identification and monosomic analysis of tan spot resistance genes in synthetic wheat lines (*Triticum turgidum* L. X *Aegilops tauschii* Coss.). Crop Sci 46: 1212–1217.

Tadesse, W., S.L.K. Hsam, and F.J. Zeller. 2006b. Evaluation of common wheat cultivars for tan spot resistance and chromosomal location of a resistance gene in the cultivar "Salamouni." Plant Breed 125(4): 318–322.

Tadesse, W., M. Schmolke, S.L.K. Hsam, V. Mohler, G. Wenzel, and F.J. Zeller. 2010. Chromosomal location and molecular mapping of a tan spot resistance gene in the winter wheat cultivar Red Chief. J Appl Genet 51(3): 235–42.

Terras, F., H. Schoofs, K. Thevissen, R.W. Osborn, J. Vanderleyden, B.P.A. Cammue, et al. 1993. Synergistic Enhancement of the Antifungal Activity of Wheat and Barley Thionins by Radish and Oilseed Rape 2S Albumins and by Barley Trypsin Inhibitors. Plant Physiol 103: 1311–1319.

Theodoulou, F.L., M. Holdsworth, and A. Baker. 2006. Peroxisomal ABC transporters. FEBS Lett 580: 1139–1155.

Thorpe, J.R., and J.L. Hall. 1984. Chronology and elicitation of changes in peroxidase and phenylalanine ammonia-lyase activities in wounded wheat leaves in response to inoculation by *Botrytis cinerea*. Physiol Plant Pathol 25: 363–379.

Tonnessen, B.W., P. Manosalva, J.M. Lang, M. Baraoidan, A. Bordeos, R. Mauleon, et al. 2015. Rice phenylalanine ammonia-lyase gene *OsPAL4* is associated with broad spectrum disease resistance. Plant Mol Biol 87: 273–286.

Tornero, P., V. Conejero, and P. Vera. 1997. Identification of a new pathogen-induced member of the subtilisin-like processing protease family from plants. J Biol Chem 272(22): 14412–14419.

Vatsa, P., A. Chiltz, S. Bourque, D. Wendehenne, A. Garcia-Brugger, and A. Pugin. 2011. Involvement of putative glutamate receptors in plant defence signaling and NO production.

Biochimie 93: 2095–2101.

Verrier, P.J., D. Bird, B. Burla, E. Dassa, C. Forestier, M. Geisler, et al. 2008. Plant ABC proteins - a unified nomenclature and updated inventory. Trends Plant Sci 13(4): 151–159.

Wang, Y., and B. Fristensky. 2001. Transgenic canola lines expressing pea defense gene DRR206 have resistance to aggressive blackleg isolates and to *Rhizoctonia solani*. Mol Breed 8: 263–271.

Wang, D., Y. Guo, C. Wu, G. Yang, Y. Li, and C. Zheng. 2008. Genome-wide analysis of CCCH zinc finger family in Arabidopsis and rice. BMC Genomics 9(44).

Wang, Y., G. Nowak, D. Culley, L. a Hadwiger, and B. Fristensky. 1999. Constitutive expression of pea defense gene DRR206 confers resistance to blackleg (*Leptosphaeria maculans*) disease in transgenic canola (*Brassica napus*). Mol Plant-Microbe Interact 12(5): 410–418.

Wang, D., K. Pajerowska-Mukhtar, A.H. Culler, and X. Dong. 2007. Salicylic Acid Inhibits Pathogen Growth in Plants through Repression of the Auxin Signaling Pathway. Curr Biol 17: 1784–1790.

Wang, S.Y., J.H. Wu, T.B. Ng, X.Y. Ye, and P.F. Rao. 2004. A non-specific lipid transfer protein with antifungal and antibacterial activities from the mung bean. Peptides 25: 1235–1242.

Wanke, D., and H. Üner Kolukisaoglu. 2010. An update on the ABCC transporter family in plants: Many genes, many proteins, but how many functions? Plant Biol 12: 15–25.

Weiberg, A., M. Wang, M. Bellinger, and H. Jin. 2014. Small RNAs: A New Paradigm in Plant-Microbe Interactions. Annu Rev Phytopathol 52: 495–516.

Wicki, W., M. Winzeler, J.E. Schmid, P. Stamp, and M. Messmer. 1999. Inheritance of resistance to leaf and glume blotch caused by Septoria nodorum Berk. in winter wheat. Theor Appl Genet 99: 1265–1272.

Williams, S.J., P. Sornaraj, E. deCourcy-Ireland, R.I. Menz, B. Kobe, J.G. Ellis, et al. 2011. An autoactive mutant of the M flax rust resistance protein has a preference for binding ATP, whereas wild-type M protein binds ADP. Mol Plant Microbe Interact 24(8): 897–906.

Williamson, V.M., V. Thomas, H. Ferris, and J. Dubcovsky. 2013. An Translocation Confers Resistance Against Root-knot Nematodes to Common Wheat. Crop Sci 53(4): 1412.

Wissenschafts, B., B. Issn, and R. October. 2000. Antifungal Activity of a Bowman-Birk-type Trypsin Inhibitor from Wheat Kernel. J Phytopathol 148: 477–481.

De Wolf, E.D., R.J. Effertz, S. Ali, and L.J. Francl. 1998. Vistas of tan spot research. Can J Plant Pathol 20(4): 349–370.

Xie, Z., L.K. Johansen, A.M. Gustafson, K.D. Kasschau, A.D. Lellis, D. Zilberman, et al. 2004. Genetic and functional diversification of small RNA pathways in plants. PLoS Biol 2(5): 642–652.

Yang, C.-W., R. Gonzalez-Lamothe, R.A. Ewan, O. Rowland, H. Yoshioka, M. Shenton, et al. 2006. The E3 Ubiquitin Ligase Activity of *Arabidopsis* PLANT U-BOX17 and Its Functional Tobacco Homolog ACRE276 Are Required for Cell Death and Defense. Plant Cell 18: 1084–1098.

Ye, X.Y., T.B. Ng, and P.F. Rao. 2001. A Bowman–Birk-Type Trypsin-Chymotrypsin Inhibitor from Broad Beans. Biochem Biophys Res Commun 289: 91–96.

Yeh, Y.-H., Y.-H. Chang, P.-Y. Huang, J.-B. Huang, and L. Zimmerli. 2015. Enhanced Arabidopsis pattern-triggered immunity by overexpression of cysteine-rich receptor-like kinases. Front Plant Sci 6: 322.

Yu, L.X., H. Barbier, M.N. Rouse, S. Singh, R.P. Singh, S. Bhavani, et al. 2014. A consensus map for Ug99 stem rust resistance loci in wheat. Theor Appl Genet 127: 1561–1581.

Yu, J., and E.S. Buckler. 2006. Genetic association mapping and genome organization of maize. Curr Opin Biotechnol 17: 155–160.

Yu, L.X., A. Lorenz, J. Rutkoski, R.P. Singh, S. Bhavani, J. Huerta-Espino, et al. 2011. Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. Theor Appl Genet 123: 1257–1268.

Yu, J., G. Pressoir, W.H. Briggs, I. Vroh Bi, M. Yamasaki, J.F. Doebley, et al. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38(2): 203–208.

Zeng, L.R., S. Qu, A. Bordeos, C. Yang, M. Baraoidan, H. Yan, et al. 2004. Spotted leaf11, A Negative Regulator of Plant Cell Death and Defense, Encodes a U-Box/Armadillo Repeat Protein Endowed with E3 Ubiquitin Ligase Activity. Plant Cell 16: 2795–2808.

Zhang, Z., E. Ersoz, C.-Q. Lai, R.J. Todhunter, H.K. Tiwari, M.A. Gore, et al. 2010. Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42(4): 355–360.

Zhang, Z., T.L. Friesen, S.S. Xu, G. Shi, Z. Liu, J.B. Rasmussen, et al. 2011. Two putatively homoeologous wheat genes mediate recognition of SnTox3 to confer effector-triggered susceptibility to *Stagonospora nodorum*. Plant J 65: 27–38.

Table S3.1: Markers significantly associated with seedling resistance to leaf rust

Marker	FDR-adjusted p-value	R^2	*T. aestivum* gene	Gene location
S3_6957300	4.91E-06	0.19	Traes_1DS_3CC12E215	6,962,963-6,970,722
			Traes_1DS_A8BD91E4A	7,031,869-7,045,046
S3_1241625	2.11E-05	0.18	Traes_1DS_3C6EAAFFD	1,236,116-1,237,549
S16_199359368	3.33E-04	0.17	Traes_6AL_5A3E5FBBD	199,359,901-199,366,693
S16_50275005	2.71E-03	0.15	Traes_6AS_EB7270F83	50,272,280-50,275,549
S10_147185899	7.24E-03	0.15	Traes_4AL_5EC714CAD	147,182,703-147,186,548
S8_40178495	1.05E-02	0.14	TRAES3BF078500390CFD_g	40,077,031-40,080,038
S16_197872823	1.08E-02	0.14	Traes_6AL_C41FC1A58	197,866,768-197,875,345
S8_13948258	1.08E-02	0.14	TRAES3BF060400070CFD_g	13,944,796-13,952,647
S8_1092429	6.09E-02	0.12	TRAES3BF035300120CFD_g	1,087,451-1,090,526
S8_667573277	6.42E-02	0.12	TRAES3BF068900010CFD_g	667,564,781-667,574,259
S5_344241063	6.44E-02	0.12	Traes_2BL_48E8EC589	344,321,397-344,325,034
S4_944423	7.18E-02	0.11	Traes_2AS_F19BE023F	941,006-945,372
S5_179485985	8.95E-02	0.09	Traes_2BL_D642EDA44	179,482,916-179,487,065
			Traes_2BL_C161B1324	179,681,358-179,689,191
S21_3991138	9.52E-02	0.09	Traes_7DS_91CCCF6DB	3,997,894-3,999,587
			Traes_7DS_1EC20E600	3,965,178-3,974,976
S5_340751784	1.66E-01	0.05	Traes_2BL_72DD69CAB	340,750,501-340,757,019
S7_8871810	1.79E-01	0.05	Traes_3AS_CBCC60379	8,867,084-8,867,737
			Traes_3AS_DFCB12564	8,949,197-8,954,357

Table S3.2: Markers significantly associated with seedling resistance to Stagonospora nodorum blotch

Marker	FDR-adjusted p-value	R^2	*T. aestivum* gene	Gene location
S7_3950435	3.28E-02	0.07	Traes_3AS_08B918CFD	3,947,040-3,948,466
S16_339241	5.00E-02	0.07	Traes_6AS_8684E560A	268,711-274,638
S5_332672189	8.78E-02	0.06	Traes_2BL_6B925F45E	332,678,123-332,681,219
S6_36849900	8.78E-02	0.06	Traes_2DL_85AFEAF70	36,849,657-36,850,483
S5_288066166	1.15E-01	0.06	Traes_2BL_1F6C61302	288,075,979-288,078,542

S4_944423	1.29E-01	0.05	Traes_2AS_F19BE023F	941,006-945,372
S19_163864204	1.29E-01	0.05	Traes_7AL_A647529F4	163,739,120-163,744,376
S4_508877	1.29E-01	0.05	Traes_2AS_6A15EE669	465,926-467,722
S14_265087281	1.29E-01	0.05	Traes_5BL_D8830A0DF	265,086,813-265,091,980
S16_201351017	1.38E-01	0.05	Traes_6AL_671AFD8EF	201,347,630-201,357,772
			Traes_6AL_EF774ECF2	201,443,118-201,445,877

Table S3.3: Markers significantly associated with seedling resistance to tan spot

Marker	FDR-adjusted p-value	R^2	*T. aestivum* gene	Gene location
S1_3589926	8.42E-02	0.1	Traes_1AS_BF353B963	3,421,247-3,425,271
			Traes_1AS_F098402B4	3,592,922-3,593,878
S7_182028651	1.36E-01	0.08	Traes_3AL_91749D67D	182,074,241-182,086,701
S4_239686345	1.36E-01	0.08	Traes_2AL_34A3B95BE	239,659,644-239,660,447
				239,659,644-239,660,447
			Traes_2AL_97FC5264A	239,704,673-239,706,147
S8_12198705	1.36E-01	0.08	TRAES3BF270500020CFD_g	12,178,410-12,179,420
S8_13415415	1.36E-01	0.07	TRAES3BF060400190CFD_g	13,425,488-13,429,397
S16_4196814	1.36E-01	0.07	Traes_6AS_3A682BA20	4,265,707-4,266,195
S8_7801088	1.36E-01	0.07	TRAES3BF060200040CFD_g	7,710,250-7,715,050
			TRAES3BF060200010CFD_g	7,797,083-7,801,241
S8_1092429	1.76E-01	0.06	TRAES3BF035300120CFD_g	1,087,451-1,090,526
S7_4804454	1.76E-01	0.06	Traes_3AS_769E90DDD	4,631,516-4,633,193
S16_9839264	1.76E-01	0.06	Traes_6AS_30C919428	9,834,233-9,840,062
S16_191519837	2.03E-01	0.05	Traes_6AL_4815187D4	191,517,427-191,519,087
S1_2331617	2.31E-01	0.05	Traes_1AS_AAB89883E1	2,295,827-2,298,821

			Traes_1AS_459048879	2,191,166-2,194,102
			Traes_1AS_3BE2A2127	2,416,572-2,418,737
S7_4563676	2.31E-01	0.05	Traes_3AS_817ECEF75	4,494,005-4,494,998
S5_281016023	2.52E-01	0.05	Traes_2BL_7C2F474DE	281,027,230-281,028,685
			Traes_2BL_D055B271C	281,011,689-281,012,413
S1_2584791	2.58E-01	0.05	Traes_1AS_C8A8A4118	2,583,319-2,585,608
			Traes_1AS_B716E0B0E	2,564,929-2,568,129
			Traes_1AS_8D33AB43B	2,555,887-2,561,122
			Traes_1AS_C7A8188D1	2,521,366-2,527,020

Table S3.4: Markers significantly associated with seedling and adult plant resistance to stripe rust

Dataset	Marker	FDR-adjusted p-value	Marker R^2	*T. aestivum* gene	Gene location
Njoro 2011	S4_208035	1.67E-01	0.09	Traes_2AS_5EB59FFC0	207,328-208,813
Quito 2012		1.03E-01	0.09		
Seedling		9.04E-07	0.17		
Toluca 2011		3.38E-05	0.14		
Toluca 2013		1.40E-05	0.16		
Njoro 2011	S4_508877	1.67E-01	0.09	Traes_2AS_6A15EE669	465,926-467,722
Quito 2012		1.27E-04	0.14		
Seedling		6.69E-11	0.27		
Toluca 2011		1.35E-04	0.13		
Toluca 2013		2.15E-07	0.2		
Njoro 2011	S4_944423	2.52E-02	0.1	Traes_2AS_F19BE023F	941,006-945,372
Quito 2012		2.63E-04	0.11		
Seedling		5.18E-10	0.24		
Toluca 2011		1.21E-04	0.13		

Toluca 2013		9.69E-08	0.2		
Quito 2012	S4_5007061	3.20E-01	0.07	Traes_2AS_6BC67DD45	5,005,258-5,009,619
Seedling		5.48E-05	0.17		
Njoro 2011	S4_5287800	3.28E-02	0.1	Traes_2AS_A477CDA77	5,203,111-5,204,650
Quito 2012		1.67E-04	0.12		
Seedling		2.46E-08	0.21		
Toluca 2011		4.04E-04	0.11		
Toluca 2013		1.39E-05	0.16		
Njoro 2011	S4_7117805	2.65E-02	0.1	Traes_2AS_6CE6AB560	7,118,296-7,123,325
Quito 2012		1.26E-03	0.1		
Seedling		1.44E-08	0.21		
Toluca 2011		2.45E-03	0.11		
Toluca 2013		4.75E-06	0.16		
Njoro 2011	S6_132714407	2.55E-01	0.05	Traes_2DL_4B5D621C1	132,713,603-132,717,046
Quito 2012	S8_17773150	6.46E-03	0.1	TRAES3BF050800140CFD_g	17,778,840-17,783,220
Toluca 2011	S8_566227604	1.75E-01	0.06	TRAES3BF027700080CFD_g	566,221,006-566,227,345
Toluca 2013	S21_4853558	7.81E-02	0.07	Traes_7DS_600B0996B	4,850,870-4,853,045

Table S3.5: Genes in the 2AS distal chromosomal region (0 to 7,123,325 bp)

Gene	Location	Orthologue	Predicted function	Species	Identity	Domain(s) in *T. aestivum* gene transcript
Traes_2AS_15D7300B6	11,485-15,872	TRIUR3_11874	Protein EIN4	*Triticum urartu*	96	CheY-like superfamily; GAF domain; Histidine kinase-like ATPase, C-terminal domain; Signal transduction histidine kinase EnvZ-like, dimerisation/phosphoacceptor domain; Signal transduction histidine kinase, hybrid-type, ethylene sensor; Signal

						transduction response regulator, receiver domain
Traes_2AS_BB7AAC01E	56,682-61,859		FMN-linked oxidoreductases superfamily protein			
Traes_2AS_88D6A69DD	78,893-79,492	F775_15220	Cytochrome P450 76C1	*Aegilops tauschii*	78	Cytochrome P450
Traes_2AS_E1DD8FB4D	194,022-195,764	TRIUR3_10292	Obtusifoliol 14-alpha demethylase	*Triticum urartu*	86	Cytochrome P450
Traes_2AS_821E10536	204,922-206,136	F775_21407	Cytochrome P450 86A1	*Aegilops tauschii*	91	Cytochrome P450
Traes_2AS_5EB59FFC0	207,328-208,813	F775_06675	Phenylalanine ammonia-lyase	*Aegilops tauschii*	92	Aromatic amino acid lyase
Traes_2AS_48D41FC99	232,316-233,553	AT4G26490	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	*Arabidopsis thaliana*	38	Late embryogenesis abundant protein, LEA-14
Traes_2AS_174193654	281,526-284,506	ONIVA08G18540	Peptidase S24/S26A/S26B/S26C family protein	*Oryza nivara*	45	Peptidase S24/S26, beta-ribbon domain
Traes_2AS_ADA59BD4F	349,942-355,345		hexokinase 3			
Traes_2AS_EA89CC3DF	356,963-372,324	GSMUA_Achr9G08570_001	histone-lysine N-methyltransferase ASHH2	*Musa acuminata*	30	AWS domain, SET domain, Zinc finger, CW-type
Traes_2AS_FC98C7EB9	430,710-432,296	GSMUA_AchrUn_randomG05100_001	Cytochrome P450 89A2	*Musa acuminata*	33	Cytochrome P450
Traes_2AS_6A15EE669	465,926-467,722	LOC100842644	ABC transporter B family member 4-like	*Brachypodium distachyon*	100	ABC transporter type 1, transmembrane domain, P-loop containing nucleoside triphosphate hydrolase, AAA+ ATPase domain
Traes_2AS_71B82606A	506,701-509,179	GSMUA_Achr8G19340_001	Membrane related protein CP5	*M. acuminata*	66	START domain
Traes_2AS_F485758C1	510,835-513,787		pyruvate dehydrogenase E1 alpha			
Traes_2AS_D917848B7	680,822-686,162		D111/G-patch domain-containing protein			
Traes_2AS_33FC65871	685,917-691,199		Eukaryotic initiation factor 3 gamma subunit family protein			

Traes_2AS_D2D7342B3	723,415-724,755	TRIUR3_14654	Cytochrome P450 86B1	*Triticum urartu*	97	Cytochrome P450
Traes_2AS_21DD59F60	787,888-788,679	MTR_3g009500	UDP-glucosyltransferase	*Medicago truncatula*	27	UDP-glucuronosyl/UDP-glucosyltransferase
Traes_2AS_1B39D2FA8	812,955-814,381					Domain of unknown function DUF4371, Ribonuclease H-like domain
Traes_2AS_09C4D333F	893,817-895,658	F775_15920	Putative NAD(P)H-dependent oxidoreductase 1	*Aegilops tauschii*	95	Aldo/keto reductase, NADP-dependent oxidoreductase domain
Traes_2AS_FEFDC29F2	900,479-905,272					Histone deacetylase complex subunit SAP30/SAP30-like
Traes_2AS_F19BE023F	941,006-945,372	TRIUR3_09185	Putative disease resistance protein RXW24L	*Triticum urartu*	92	
Traes_2AS_C075AEE6C	1,006,381-1,007,592	TRIUR3_10294	Cytochrome P450 76C2	*Triticum urartu*	82	Cytochrome P450
Traes_2AS_048E13951	1,016,839-1,021,303					Protein phosphatase 2C
Traes_2AS_DE0A69F86	1,029,392-1,031,392	F775_15954	Cytochrome P450 86A1	*Aegilops tauschii*	54	Cytochrome P450
Traes_2AS_9A9E4CC41	1,040,393-1,040,677					HAT dimerisation domain, C-terminal, Ribonuclease H-like domain
Traes_2AS_ACFF3CB9F	1,064,508-1,065,138					NB-ARC
Traes_2AS_AFEB22E37	1,078,143-1,079,932					P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_FA7C1F225	1,125,533-1,127,020	TCM_015276	UDP-glucosyl transferase 88A1, putative	*Theobroma cacao*	52	UDP-glucuronosyl/UDP-glucosyltransferase
Traes_2AS_A087099AF	1,183,721-1,186,448					
Traes_2AS_D57565E23	1,186,917-1,192,080					
Traes_2AS_9F5A6C54F	1,199,816-1,201,857	OS01G0804900	Cytochrome P450-dependent fatty acid hydroxylase-like protein	*Oryza sativa Japonica*	64	Cytochrome P450
Traes_2AS_7AE2FEA16	1,211,604-1,212,557					Uncharacterised protein family UPF0546
Traes_2AS_E82763741	1,283,975-1,284,799	GSMUA_Achr2G10160_001	Putative Cytochrome P450 86B1	*Musa acuminata*	32	Cytochrome P450

Traes_2AS_5BA299224	1,306,297-1,309,224					Major facilitator superfamily domain, Proton-dependent oligopeptide transporter family
Traes_2AS_6EB8AD265	1,345,642-1,348,508					Leucine-rich repeat domain, L domain-like
Traes_2AS_6DEE8FF80	1,386,842-1,388,352	TRIUR3_07758	Cytochrome P450 86B1	*Triticum urartu*	84	Cytochrome P450
Traes_2AS_D2DB0E357	1,413,392-1,425,005	GSMUA_Achr4G29640_001	Polycomb group protein EMBRYONIC FLOWER 2	*Musa acuminata*	32	Polycomb protein, VEFS-Box
Traes_2AS_AC1EE71AB	1,471,456-1,472,041					Protein of unknown function DUF594
Traes_2AS_7581E280A	1,506,343-1,511,761	TRIUR3_01916	ABC transporter B family member 11	*Triticum urartu*	90	ABC transporter type 1, transmembrane domain, P-loop containing nucleoside triphosphate hydrolase, AAA+ ATPase domain
Traes_2AS_00F345CBB	1,515,753-1,517,096					
Traes_2AS_6F531FBB5	1,524,389-1,528,207f					ARID DNA-binding domain, ELM2 domain
Traes_2AS_5CAF7A3671	1,608,937-1,610,457	TRIUR3_10228	Putative disease resistance protein RGA3	*Triticum urartu*	99	NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_0898755211	1,613,015-1,614,568	TRIUR3_10229	Putative disease resistance RPP13-like protein 1	*Triticum urartu*	100	P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_6E1B0E6EF	1,619,489-1,624,160	TRIUR3_11002	Disease resistance protein RGA2	*Triticum urartu*	86	P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_5A921E630	1,642,311-1,645,309	TRIUR3_00568	Sec23/Sec24 protein transport family protein	*Triticum urartu*	90	ADF-H/Gelsolin-like domain, Gelsolin-like domain, Sec23/Sec24 beta-sandwich
Traes_2AS_6BB5BCCA2	1,668,886-1,673,001	TRIUR3_00558	Werner Syndrome-like exonuclease	*Triticum urartu*	88	Ribonuclease H-like domain
Traes_2AS_E1C66CA12	1,673,131-1,676,858					Cytochrome oxidase assembly protein 1
Traes_2AS_5BE64EB9E	1,768,877-1,769,792					Cytochrome P450
Traes_2AS_D66FB2E90	1,775,306-1,775,420					Transcription factor TGA like domain
Traes_2AS_5A9348E72	1,836,189-1,840,014					
Traes_2AS_B5D438C6C	1,852,098-1,855,412					Protein of unknown function DUF724, Tudor-like, plant

Traes_2AS_A8793AD4F	1,873,413-1,878,909	TCM_006167	Yth domain-containing protein, putative isoform 1	*Theobroma cacao*	26	YTH domain
Traes_2AS_E82A7163B	1,880,594-1,884,410					Protein of unknown function DUF1644, Zinc finger, RING/FYVE/PHD-type
Traes_2AS_1E8BDB2C6	1,896,429-1,902,311	TRIUR3_01918	ABC transporter B family member 4	*Triticum urartu*	99	ABC transporter type 1, transmembrane domain, P-loop containing nucleoside triphosphate hydrolase, AAA+ ATPase domain
Traes_2AS_B264257CD	1,916,608-1,917,879	TRIUR3_00556	Flavin-containing monooxygenase YUCCA8	*Triticum urartu*	100	Pyridine nucleotide-disulphide oxidoreductase, FAD/NAD(P)-binding domain
Traes_2AS_D0C21ADB5	1,937,418-1,938,543					DNA-binding WRKY
Traes_2AS_5CAF7A367	1,977,217-1,978,737	TRIUR3_10228	Putative disease resistance protein RGA3	*Triticum urartu*	99	NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_089875521	1,981,295-1,982,848	TRIUR3_10229	Putative disease resistance RPP13-like protein 1	*Triticum urartu*	100	P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_EFB7DF837	1,987,769-1,992,440	TRIUR3_11002	Disease resistance protein RGA2	*Triticum urartu*	87	P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_5C5EAEAB8	2,051,212-2,053,364	TRIUR3_01915	Tyrosine N-monooxygenase	*Triticum urartu*	98	Cytochrome P450
Traes_2AS_B096B5D19	2,059,096-2,059,334					
Traes_2AS_9ED98D871	2,114,563-2,115,975	BRADI5G01280	Sec23/Sec24 protein transport family protein	*Brachypodium distachyon*	64	Sec23/Sec24, trunk domain, von Willebrand factor, type A
Traes_2AS_7B74BE6D6	2,179,779-2,183,218	F775_25605	Putative LRR receptor-like serine/threonine-protein kinase	*Aegilops tauschii*	96	Concanavalin A-like lectin/glucanase domain, Leucine-rich repeat, Serine/threonine-protein kinase
Traes_2AS_0C6E89B86	2,185,152-2,193,592		UDP-Glycosyltransferase superfamily protein			UDP-Glycosyltransferase
Traes_2AS_6E3F18D20	2,328,227-2,333,933	BRADI5G01220	AAR2 protein family	*Brachypodium distachyon*	90	A1 cistron-splicing factor, AAR2
Traes_2AS_1171AE3E3	2,368,039-2,368,281					
Traes_2AS_90E3B2E76	2,371,741-2,378,409	TRIUR3_00568	Sec23/Sec24 protein transport family protein	*Triticum urartu*	98	ADF-H/Gelsolin-like domain, Gelsolin-like domain, Sec23/Sec24 beta-sandwich, Sec23/Sec24, helical domain,

						Sec23/Sec24, trunk domain, Zinc finger, Sec23/Sec24-type, von Willebrand factor, type A
Traes_2AS_D38E1E530	2,425,561-2,429,488	TRIUR3_33851	Cytochrome P450 85A1	*Triticum urartu*	97	Cytochrome P450
Traes_2AS_1493D0ADB	2,434,789-2,438,797	TRIUR3_16698	TOM1-like protein 2	*Triticum urartu*	100	ENTH/VHS, GAT domain, VHS domain, Target of Myb protein 1
Traes_2AS_D2FB48992	2,511,299-2,515,640	TRIUR3_01922	Arginine/serine-rich-splicing factor RSP31	*Triticum urartu*	97	Nucleotide-binding alpha-beta plait domain, RNA recognition motif domain
Traes_2AS_D3715AC4A	2,600,647-2,602,717	BRADI2G40400	alpha/beta-Hydrolases superfamily protein	*Brachypodium distachyon*	71	Alpha/Beta hydrolase fold, Dienelactone hydrolase
Traes_2AS_48DA475B0	2,605,114-2,608,922	POPTR_0008s20920	NAD(P)-binding Rossmann-fold superfamily protein	*Populus trichocarpa*	68	NAD(P)-binding domain
Traes_2AS_1FB8F6760	2,610,953-2,611,098	F775_30926	Nicotianamine synthase-like 5 protein	*Aegilops tauschii*	100	Nicotianamine synthase
Traes_2AS_AFA56788E	2,631,651-2,635,406	OS03G0602300	Cytochrome P450 85A1	*Oryza sativa Japonica*	84	Cytochrome P450
Traes_2AS_C8325AEE0	2,672,752-2,673,312					
Traes_2AS_0E7069178	2,682,863-2,688,664	TRIUR3_15844	E3 ubiquitin ligase BIG BROTHER-related protein	*Triticum urartu*	99	Zinc finger, RING-type
Traes_2AS_F8779F643	2,708,791-2,716,281					Serine/threonine-protein kinase
Traes_2AS_27D2CB02A	2,735,584-2,743,313	TRIUR3_03051	Mortality factor 4-like protein 1	*Triticum urartu*	71	Chromo domain-like, MRG domain
Traes_2AS_18956EE53	2,744,835-2,746,252	GSMUA_Achr4G26690_001	pectate lyase 15	*Musa acuminata*	85	Pectate lyase, AmbAllergen
Traes_2AS_0C6583718	2,746,355-2,748,906	GRMZM2G157354	Nicotiana lesion-inducing like protein	*Zea mays*	57	HR-like lesion-inducer
Traes_2AS_D35789679	2,781,055-2,784,866	TCM_001771	4-hydroxyphenylacetaldehyde oxime monooxygenase, putative	*Theobroma cacao*	51	Cytochrome P450
Traes_2AS_7E07335F8	2,809,078-2,810,795	F775_17719	Tyrosine N-monooxygenase	*Aegilops tauschii*	94	Cytochrome P450
Traes_2AS_1E76C16FF	2,824,218-2,826,100					P-loop containing nucleoside triphosphate hydrolase

Traes_2AS_ECCD1E063	2,827,522-2,831,276	GSMUA_Achr3G09670_001	ABC transporter B family member	*Musa acuminata*	65	ABC transporter type 1, transmembrane domain
Traes_2AS_D7F13D5C9	2,831,827-2,833,805	POPTR_0001s29260	fatty acid amide hydrolase	*Populus trichocarpa*	49	Amidase signature domain
Traes_2AS_BBE5262D9	2,838,542-2,840,321					
Traes_2AS_A8BF6F543	2,845,738-2,847,421		heat stable protein 1			Dimeric alpha-beta barrel, Stress responsive alpha-beta barrel
Traes_2AS_C8DB2D8D6	2,857,320-2,859,807	F775_01098	Cysteine-rich receptor-like protein kinase 10	*Aegilops tauschii*	82	Serine/threonine-protein kinase
Traes_2AS_802012621	2,876,318-2,878,062	TRIUR3_22387	Wall-associated receptor kinase 3	*Triticum urartu*	95	Concanavalin A-like lectin/glucanase domain, EGF-like calcium-binding domain, Insulin-like growth factor binding protein, N-terminal, Serine/threonine-protein kinase
Traes_2AS_5907AC060	2,902,811-2,903,319					
Traes_2AS_55A5C5791	2,920,103-2,920,895	Bo9g059980	FAD-binding Berberine family protein	*Brassica oleraceae*	43	Berberine/berberine-like
Traes_2AS_55FB06D16	2,929,728-2,930,679	GRMZM2G124785	Nicotianamine synthase 2	*Zea mays*	41	Nicotianamine synthase
Traes_2AS_A3F743F75	2,942,903-2,943,796					
Traes_2AS_2E21BFFFB	2,951,179-2,953,372	F775_17990	Chalcone synthase	*Aegilops tauschii*	94	Chalcone/stilbene synthase, Polyketide synthase, type III, Thiolase like
Traes_2AS_57DFD533D	2,983,516-2,984,167					
Traes_2AS_8770C8C36	3,003,227-3,003,474					
Traes_2AS_6C89A3064	3,014,100-3,020,389		ELMO/CED-12 family protein			Engulfment/cell motility, ELMO
Traes_2AS_08A91D578	3,021,419-3,021,957					
Traes_2AS_B96BB2490	3,031,360-3,038,368		zinc finger (CCCH-type) family protein			(Uracil-5)-methyltransferase family, Nucleotide-binding alpha-beta plait domain, RNA methyltransferase TrmA, active site, S-adenosyl-L-methionine-dependent methyltransferase

Traes_2AS_F2A8F88A3	3,062,168-3,063,407	OS11G0432900	Serine carboxypeptidase family protein	*Oryza sativa Japonica*	76	Alpha/Beta hydrolase fold, Peptidase S10, serine carboxypeptidase
Traes_2AS_512E437B5	3,070,839-3,071,330	F775_14818	Cytochrome P450 71C2	*Aegilops tauschii*	98	Cytochrome P450
Traes_2AS_A3B0031A6	3,072,133-3,073,639	F775_14818	Cytochrome P450 71C2	*Aegilops tauschii*	96	Cytochrome P450
Traes_2AS_AC62A3F1C	3,080,626-3,083,718					Domain of unknown function DUF4220
Traes_2AS_7D25F69CB	3,132,390-3,135,807	PRUPE_ppa008443mg	Nucleotide-diphospho-sugar transferase family protein	*Prunus persica*	22	Nucleotide-diphospho-sugar transferase
Traes_2AS_FD9DC88CA1	3,147,755-3,150,987	BRADI5G02990	FKBP-like peptidyl-prolyl cis-trans isomerase family protein	*Brachypodium distachyon*	73	Peptidyl-prolyl cis-trans isomerase, FKBP-type
Traes_2AS_0B776BDEE	3,161,720-3,162,254					Cytochrome P450
Traes_2AS_592926586	3,247,963-3,249,013					D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding domain, NAD(P)-binding domain
Traes_2AS_B8091AC54	3,300,384-3,305,959	TRIUR3_23980	Disease resistance protein RPM1	*Triticum urartu*	71	Leucine-rich repeat domain, L domain-like, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_5A0BCB6E0	3,306,182-3,311,622					
Traes_2AS_32405E54F	3,314,219-3,318,690		Formamidopyrimidine-DNA glycosylase	*Aegilops tauschii*	85	DNA glycosylase/AP lyase, H2TH DNA-binding, Ribosomal protein S13-like, H2TH
Traes_2AS_A4BAAEE05	3,350,821-3,353,514	OS04G0178300	Syn-copalyl diphosphate synthase	*Oryza sativa Japonica*	65	Terpene synthase, Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid
Traes_2AS_E6FF6A652	3,360,901-3,361,678	BRADI4G17230	Chalcone and stilbene synthase family protein	*Brachypodium distachyon*	89	Chalcone/stilbene synthase, C-terminal, Thiolase like
Traes_2AS_972F2CECA	3,399,290-3,408,485	TRIUR3_02866	Cullin-associated NEDD8-dissociated protein 1	*Triticum urartu*	94	Armadillo-like helical, TATA-binding protein interacting (TIP20)
Traes_2AS_86E424632	3,417,793-3,420,233	TRIUR3_19384	Myrcene synthase, chloroplastic	*Triticum urartu*	86	Isoprenoid synthase domain, Terpene synthase, N-terminal domain, Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid

Traes_2AS_A0DB10548	3,420,922-3,423,704	GSMUA_Achr10G09080_001	CDT1A - Putative DNA replication initiation protein, expressed	*Musa acuminata*	46	CDT1 Geminin-binding domain-like, DNA replication factor Cdt1, C-terminal, Winged helix-turn-helix DNA-binding domain
Traes_2AS_AD32B0419	3,425,483-3,427,078	TRIUR3_28181	Reticuline oxidase-like protein	*Triticum urartu*	50	Berberine/berberine-like, CO dehydrogenase flavoprotein-like, FAD-binding, subdomain 2, FAD linked oxidase, N-terminal, FAD-binding, type 2
Traes_2AS_607E72B1A	3,482,676-3,484,835	F775_13064	1-aminocyclopropane-1-carboxylate oxidase-1-like protein	*Aegilops tauschii*	96	Isopenicillin N synthase-like, Non-haem dioxygenase N-terminal domain, Oxoglutarate/iron-dependent dioxygenase
Traes_2AS_0FA7292A9	3,513,535-3,528,917					
Traes_2AS_D38A5E045	3,531,826-3,536,363	F775_05049	Putative amidase	*Aegilops tauschii*	67	Amidase signature domain
Traes_2AS_E5D366E76	3,542,721-3,544,462					Cytochrome P450
Traes_2AS_03158984F	3,621,705-3,622,356					
Traes_2AS_737B8B16E	3,680,868-3,681,101					
Traes_2AS_195F8322C	3,696,207-3,697,724					
Traes_2AS_8588747E6	3,721,533-3,724,798					
Traes_2AS_292D0991C	3,764,073-3,765,666					Major facilitator superfamily domain
Traes_2AS_0CABCA673	3,780,965-3,785,081	F775_20787	ABC transporter C family member 4	*Aegilops tauschii*	82	ABC transporter type 1, transmembrane domain, P-loop containing nucleoside triphosphate hydrolase, AAA+ ATPase domain
Traes_2AS_F15D136CF	3,811,693-3,812,895	TRIUR3_03937	Putative glutathione S-transferase GSTF1	*Triticum urartu*	95	Glutathione S-transferase, Thioredoxin-like fold
Traes_2AS_712C217B9	3,822,679-3,823,672					
Traes_2AS_EDB52A1EB	3,825,633-3,827,560	F775_01485	Protein RUPTURED POLLEN GRAIN 1	*Aegilops tauschii*	70	SWEET sugar transporter

Traes_2AS_E3638EE97	3,962,722-3,963,023					
Traes_2AS_9D1FC09D7	3,978,834-3,982,971					Leucine-rich repeat domain, L domain-like, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_0C03C4F29	3,993,406-3,993,989	TRIUR3_06927	Defensin-like protein 2	*Triticum urartu*	93	Gamma Purothionin, Knottin, scorpion toxin-like
Traes_2AS_37A335CE4	3,999,946-4,002,582	F775_02916	Formin-like protein 11	*Aegilops tauschii*	94	Formin, FH2 domain
Traes_2AS_29CF270C8	4,055,542-4,056,267	F775_14593	Defensin-like protein P322	*Aegilops tauschii*	95	Gamma Purothionin, Knottin, scorpion toxin-like
Traes_2AS_0E33CD71F	4,131,971-4,132,441					Nucleotide-diphospho-sugar transferase
Traes_2AS_0E225338F	4,185,852-4,186,590	F775_11453	Chemocyanin	*Aegilops tauschii*	72	Cupredoxin, Plastocyanin-like
Traes_2AS_484015532	4,217,044-4,217,949					
Traes_2AS_49F9C2E5D	4,249,429-4,252,279					Peptidase C65, otubain
Traes_2AS_453C10ECC	4,273,730-4,291,014	Si009627m.g	fatty acid amide hydrolase	*Setaria italica*	82	Amidase
Traes_2AS_3F13B523F	4,324,394-4,325,943	Bo8g116100	splicing factor Prp18 family protein	*Brassica oleraceae*	78	Prp18
Traes_2AS_82FF45D92	4,360,958-4,361,268					CO dehydrogenase flavoprotein-like, FAD-binding, subdomain 2, FAD linked oxidase, N-terminal, FAD-binding, type 2
Traes_2AS_C913A8043	4,470,322-4,475,786	TRIUR3_14868	G-type lectin S-receptor-like serine/threonine-protein kinase	*Triticum urartu*	99	Bulb-type lectin domain, Concanavalin A-like lectin/glucanase domain, Serine/threonine-protein kinase,
Traes_2AS_1D2CDFD90	4,490,058-4,490,782					
Traes_2AS_D61611BDC	4,539,695-4,540,797					
Traes_2AS_D537D01E3	4,602,759-4,603,439					
Traes_2AS_AA9C737C7	4,632,431-4,634,488	F775_13064	1-aminocyclopropane-1-carboxylate oxidase-1-like protein	*Aegilops tauschii*	93	Isopenicillin N synthase-like, Oxoglutarate/iron-dependent dioxygenase

Traes_2AS_786CFD895	4,645,645-4,646,384					Glycoside hydrolase
Traes_2AS_1A307CE26	4,656,211-4,656,893	OS08G0432300	Putative teosinte branched1 protein	*Oryza sativa Japonica*	53	Transcription factor TCP subgroup
Traes_2AS_768355513	4,674,801-4,676,429	TRIUR3_13099	1-aminocyclopropane-1-carboxylate oxidase-like protein 2	*Triticum urartu*	62	Isopenicillin N synthase-like, Non-haem dioxygenase N-terminal domain, Oxoglutarate/iron-dependent dioxygenase
Traes_2AS_4FEC7A465	4,699,741-4,704,668	F775_17871	Syn-copalyl diphosphate synthase	*Aegilops tauschii*	74	Isoprenoid synthase domain, Terpene synthase, Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid
Traes_2AS_04D7D387C	4,776,276-4,777,628					
Traes_2AS_CAA2FBF93	4,882,731-4,886,348	TRIUR3_08998	Beta-glucosidase 16	*Triticum urartu*	88	Glycoside hydrolase
Traes_2AS_67817FEBC	4,891,171-4,894,255	GSMUA_Achr1G09280_001	serine-type peptidase	*Musa acuminata*	63	Peptidase
Traes_2AS_6D59C67F0	4,900,479-4,915,398	GRMZM2G154509	Phosphoglycerate mutase-like protein isoform 1	*Zea mays*	72	Histidine phosphatase superfamily
Traes_2AS_AD1248C31	4,921,435-4,921,732					Chloramphenicol acetyltransferase-like domain
Traes_2AS_2EB281841	4,937,584-4,941,371	OS10G0124500	F-box domain containing protein	*Oryza sativa Japonica*	51	F-box domain
Traes_2AS_09E707C65	4,948,052-4,951,281	TRIUR3_12520	Cysteine-rich receptor-like protein kinase 41	*Triticum urartu*	82	Concanavalin A-like lectin/glucanase domain, MSP domain, PapD-like, Protein kinase
Traes_2AS_E0AFA6D18	4,972,252-4,975,044	F775_14936	Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 2	*Aegilops tauschii*	96	ATP-sulfurylase PUA-like domain, PUA-like domain, Rossmann-like alpha/beta/alpha sandwich fold, Sulphate adenylyltransferase
Traes_2AS_6BC67DD45	5,005,258-5,009,619	TRIUR3_16539	Putative disease resistance RPP13-like protein 1	*Triticum urartu*	97	Leucine-rich repeat domain, L domain-like
Traes_2AS_FBBAC0883	5,092,563-5,096,155	GSMUA_Achr2G11220_001	Putative amidase	*Musa acuminata*	61	Amidase
Traes_2AS_82A750758	5,104,285-5,107,870	TRIUR3_01665	Putative acetyl-CoA acetyltransferase, cytosolic 2	*Triticum urartu*	91	Thiolase-like
Traes_2AS_A9F768C2B	5,111,270-5,112,120	GRMZM2G068947	12-oxo-phytodienoic acid reductase	*Zea mays*	44	Aldolase-type TIM barrel, NADH:flavin oxidoreductase/NADH oxidase, N-terminal

Traes_2AS_52D58DF7F	5,112,603-5,114,095	TCM_046775	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	*Theobroma cacao*	57	Isopenicillin N synthase-like, Oxoglutarate/iron-dependent dioxygenase
Traes_2AS_74C2A5D1F	5,139,357-5,141,697	TRIUR3_20816	Disease resistance RPP13-like protein 4	*Triticum urartu*	40	Leucine-rich repeat domain, L domain-like
Traes_2AS_A477CDA77	5,203,111-5,204,650	TRIUR3_30356	Putative disease resistance RPP13-like protein 1	*Triticum urartu*	71	P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_C475CC0F9	5,284,999-5,288,091	F775_11445	Vacuolar amino acid transporter 1	*Aegilops tauschii*	98	Amino acid transporter, transmembrane
Traes_2AS_8228ECD84	5,326,512-5,327,522	MLOC_54115	peptidylprolyl cis/trans isomerase, NIMA-interacting 1	*Hordeum vulgare*	69	Peptidyl-prolyl cis-trans isomerase, PpiC-type
Traes_2AS_FD956B4EF	5,347,125-5,348,654	MTR_4g128210	muniscin carboxy-terminal mu-like domain protein	*Medicago truncatula*	55	Mu homology domain, Muniscin C-terminal
Traes_2AS_7A4976FA9	5,360,566-5,363,634	GSMUA_Achr11G02290_001	CBS domain containing membrane protein	*Musa acuminata*	51	CBS domain
Traes_2AS_FEBDEF579	5,397,058-5,397,826					
Traes_2AS_F06C61AB1	5,429,704-5,431,056					P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_38949B685	5,437,919-5,441,973	F775_06790	Serine carboxypeptidase-like 19	*Aegilops tauschii*	90	Alpha/Beta hydrolase fold, Peptidase S10, serine carboxypeptidase
Traes_2AS_03F50010A	5,469,915-5,470,507					
Traes_2AS_85E4E40C3	5,530,550-5,532,694	F775_00200	Cytokinin-O-glucosyltransferase 2	*Aegilops tauschii*	55	UDP-glucuronosyl/UDP-glucosyltransferase
Traes_2AS_54398D868	5,551,825-5,553,397					P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_28B3FAEB3	5,557,208-5,561,275	OS12G0123500	Probable apyrase 3	*Oryza sativa Japonica*	70	Nucleoside phosphatase GDA1/CD39
Traes_2AS_2821A7128	5,584,894-5,586,491	F775_15232	Anthranilate N-benzoyltransferase protein 1	*Aegilops tauschii*	72	Chloramphenicol acetyltransferase-like domain, transferse
Traes_2AS_F25B2DA46	5,620,118-5,621,330					
Traes_2AS_CED8C9B73	5,642,653-5,650,977	F775_01103	Putative histone acetyltransferase HAC-like protein 3	*Aegilops tauschii*	76	CBP/p300-type histone acetyltransferase domain, Histone H3-K56 acetyltransferase, RTT109, Zinc finger, FYVE/PHD-type, Zinc finger, TAZ-type, Zinc finger, ZZ-type

Traes_2AS_3D0A18D67	5,660,860-5,663,270					Glycosyl transferase, family 43, Nucleotide-diphospho-sugar transferases
Traes_2AS_D11AD107B	5,664,472-5,667,082	GSMUA_Achr9G10980_001	Putative Eukaryotic translation initiation factor 3 subunit M	*Musa acuminata*	64	Proteasome component (PCI) domain, Winged helix-turn-helix DNA-binding domain
Traes_2AS_BA5EB39CB	5,794,485-5,796,380	OS08G0157500	Flavone 3'-O-methyltransferase 1	*Oryza sativa Japonica*	65	O-methyltransferase COMT-type, Plant methyltransferase dimerisation, S-adenosyl-L-methionine-dependent methyltransferase, Winged helix-turn-helix DNA-binding domain
Traes_2AS_B78968A63	5,812,028-5,821,870	F775_28219	Cyclopropane-fatty-acyl-phospholipid synthase	*Aegilops tauschii*	91	Amine oxidase, Pyridine nucleotide-disulphide oxidoreductase, FAD/NAD(P)-binding domain 3.50.50.60, S-adenosyl-L-methionine-dependent methyltransferase
Traes_2AS_8D7326365	5,878,646-5,880,443	F775_12427	Protein RUPTURED POLLEN GRAIN 1	*Aegilops tauschii*	98	SWEET sugar transporter
Traes_2AS_43F1E3236	5,889,641-5,890,252					
Traes_2AS_C6EE01E7C	5,925,019-5,925,190					DNA glycosylase/AP lyase, catalytic domain
Traes_2AS_11CD58DCB	5,960,922-5,961,200					
Traes_2AS_7D58356F2	5,984,614-5,988,797	LOC100830175	Disease resistance protein RGA3	*Brachypodium distachyon*	79	Leucine-rich repeat domain, L domain-like, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_0F0720621	6,022,294-6,025,479	F775_04990	Wall-associated receptor kinase 2	*Aegilops tauschii*	92	EGF-like calcium-binding domain, EGF-like calcium-binding, conserved site, Insulin-like growth factor binding protein, N-terminal, Serine-threonine/tyrosine-protein kinase catalytic domain
Traes_2AS_892049962	6,028,623-6,030,299	GSMUA_Achr9G17450_001	Sugar transporter ERD6-like 5	*Musa acuminata*	71	General substrate transporter, Major facilitator superfamily domain
Traes_2AS_F3D9CA560	6,041,291-6,046,379	BRADI5G01960	zinc finger (Ran-binding) family protein	*Brachypodium distachyon*	79	Zinc finger, RanBP2-type
Traes_2AS_E55705603	6,047,423-6,049,703	GSMUA_Achr7G16470_001	Probable carbohydrate esterase At4g34215	*Musa acuminata*	48	SGNH hydrolase-type esterase domain

Traes_2AS_9C459F3AA	6,051,185-6,056,833	F775_26585	Pumilio-like protein	*Aegilops tauschii*	96	Armadillo-like helical, Pumilio RNA-binding repeat
Traes_2AS_84673C57E	6,115,027-6,116,935	GSMUA_Achr3G25740_001	Serine carboxypeptidase-like 18	*Musa acuminata*	45	Alpha/Beta hydrolase fold, Peptidase S10, serine carboxypeptidase
Traes_2AS_1742B494A	6,146,931-6,152,135	TRIUR3_09802	actin-related protein 9	*Triticum urartu*	75	Actin-related protein 8/Plant actin-related protein 9
Traes_2AS_E899EC1A7	6,187,228-6,189,784					Peptidase S54, rhomboid domain, Protein of unknown function DUF1751, integral membrane, eukaryotic
Traes_2AS_FD9DC88CA	6,200,795-6,204,027	BRADI5G02990	FKBP-like peptidyl-prolyl cis-trans isomerase family protein	*Brachypodium distachyon*	73	Peptidyl-prolyl cis-trans isomerase, FKBP-type
Traes_2AS_B0F02BF94	6,214,894-6,215,271	MTR_3g057990	cytochrome P450 family 71 protein	*Medicago truncatula*	29	Cytochrome P450
Traes_2AS_D0CC98BD2	6,233,916-6,234,981	GSMUA_Achr2G11220_001	Putative amidase C869.01	*Musa acuminata*	58	Amidase
Traes_2AS_B1E630276	6,264,733-6,268,265	GRMZM2G164036	Cytochrome P450 CYP71C36	*Zea mays*	57	Cytochrome P450
Traes_2AS_EF55AF9DD	6,306,786-6,312,638	GRMZM2G079538	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	*Zea mays*	88	2-oxo acid dehydrogenase, lipoyl-binding site, Biotin/lipoyl attachment, Chloramphenicol acetyltransferase-like domain, Dihydrolipoamide succinyltransferase, Single hybrid motif
Traes_2AS_3478D9D66	6,318,771-6,323,180	TRIUR3_34055	Putative disease resistance RPP13-like protein 1	*Triticum urartu*	95	Leucine-rich repeat domain, L domain-like, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_DF1AB5AC4	6,441,240-6,442,023	AT2G30830	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	*Arabidopsis thaliana*	43	Isopenicillin N synthase-like, Non-haem dioxygenase N-terminal domain
Traes_2AS_3947F19A9	6,471,670-6,474,448	TRIUR3_11745	Disease resistance protein RPP13	*Triticum urartu*	81	Leucine-rich repeat domain, L domain-like, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_7934B4038	6,491,252-6,491,568	F775_14065	Putative disease resistance protein	*Aegilops tauschii*	59	NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_5815C0679	6,547,393-6,547,841	F775_15110	Putative amidase	*Aegilops tauschii*	81	Amidase
Traes_2AS_0EF4A20E2	6,548,163-6,548,604					D-isomer specific 2-hydroxyacid dehydrogenase, catalytic domain, NAD(P)-binding domain

Traes_2AS_9DF4284D3	6,579,460-6,579,783					
Traes_2AS_707CCA800	6,587,459-6,589,138					P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_537A8D6B5	6,646,448-6,652,586	AT5G60740	ABC transporter G family member 28	*Arabidopsis thaliana*	53	AAA+ ATPase domain, ABC transporter-like, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_148383FB6	6,712,387-6,715,389	F775_03222	Putative amidase	*Aegilops tauschii*	93	Amidase
Traes_2AS_B6A9A23D7	6,755,925-6,757,154	TRIUR3_10639	1-aminocyclopropane-1-carboxylate oxidase-like protein 11	*Triticum urartu*	93	Isopenicillin N synthase-like, Oxoglutarate/iron-dependent dioxygenase
Traes_2AS_3E73A7BD8	6,799,897-6,802,578	F775_12436	Protein RUPTURED POLLEN GRAIN 1	*Aegilops tauschii*	88	*SWEET sugar transporter*
Traes_2AS_F55243E0C	6,826,532-6,827,890	F775_26609	Putative nicotianamine synthase 2	*Aegilops tauschii*	97	Nicotianamine synthase, S-adenosyl-L-methionine-dependent methyltransferase
Traes_2AS_66BFD95D4	6,832,418-6,834,861	GSMUA_Achr11G19150_001	Putative F-box domain containing protein	*Musa acuminata*	23	F-box domain, Protein of unknown function DUF295
Traes_2AS_C8B6AF996	6,849,484-6,850,930					P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_150D52D59	6,864,263-6,866,588	F775_19462	GDSL esterase/lipase	*Aegilops tauschii*	86	Lipase, GDSL, SGNH hydrolase-type esterase domain
Traes_2AS_0CA29A19C	6,882,712-6,895,168	GRMZM2G154509	Phosphoglycerate mutase-like protein isoform	*Zea mays*	73	Histidine phosphatase superfamily
Traes_2AS_866633353C	6,897,466-6,901,759	TRIUR3_13083	Disease resistance protein RPM1	*Triticum urartu*	95	Leucine-rich repeat domain, L domain-like, NB-ARC, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_DAA5EF278	6,926,483-6,930,173	TRIUR3_03939	Cytochrome P450 71C2	*Triticum urartu*	85	Cytochrome P450
Traes_2AS_600F98C3A	6,959,604-6,963,872		P-loop containing nucleoside triphosphate hydrolases superfamily protein			Adenylate kinase/UMP-CMP kinase, P-loop containing nucleoside triphosphate hydrolase
Traes_2AS_10B992D89	6,968,455-6,972,302	TCM_001771	4-hydroxyphenylacetaldehyde oxime monooxygenase	*Theobroma cacao*	51	Cytochrome P450
Traes_2AS_C1EA81EC4	7,039,670-7,043,193		RuBisCO large subunit-binding protein subunit alpha, chloroplastic			Chaperonin Cpn60, GroEL-like apical domain

Traes_2AS_AD125CB18	7,055,379-7,058,871	F775_52491	arginase	*Aegilops tauschii*	91	Ureohydrolase
Traes_2AS_CFA978965	7,083,225-7,085,187	F775_13313	Beta-glucosidase 1	*Aegilops tauschii*	90	Glycoside hydrolase superfamily
Traes_2AS_5233F6588	7,112,799-7,113,498					
Traes_2AS_6CE6AB560	7,118,296-7,123,325	LOC100830175	Putative disease resistance protein RGA3	*Brachypodium distachyon*	79	Leucine-rich repeat domain, L domain-like, NB-ARC, P-loop containing nucleoside triphosphate hydrolase

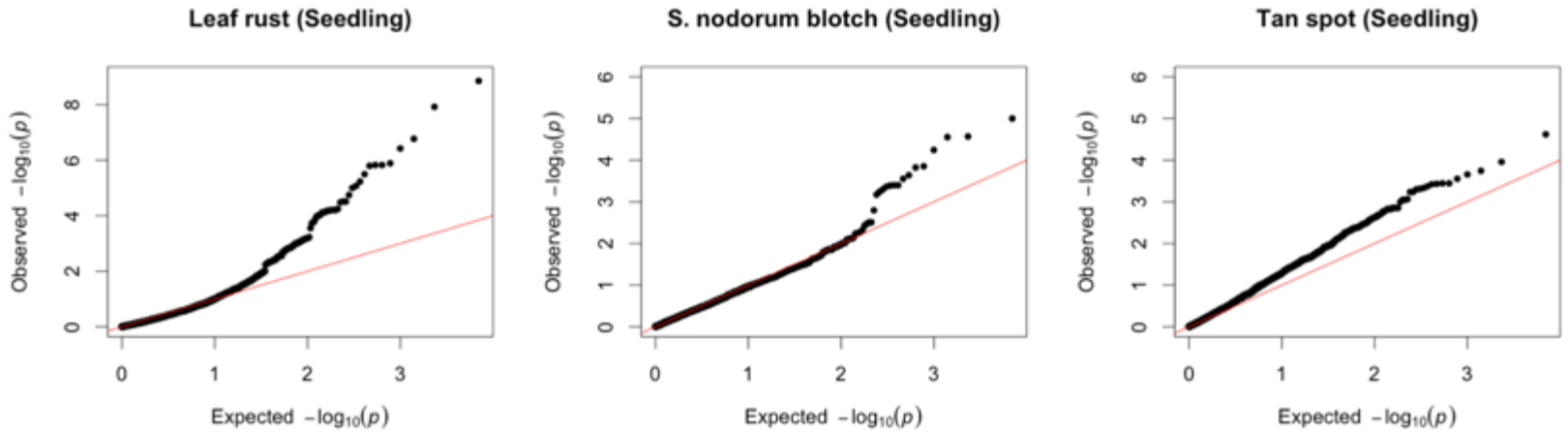

Figure S3.1: Quantile-quantile plot of the p-values comparing the p-value distribution to a uniform null distribution for seedling resistance to leaf rust, Stagonospora nodorum blotch and tan spot in the 45th IBWSN.

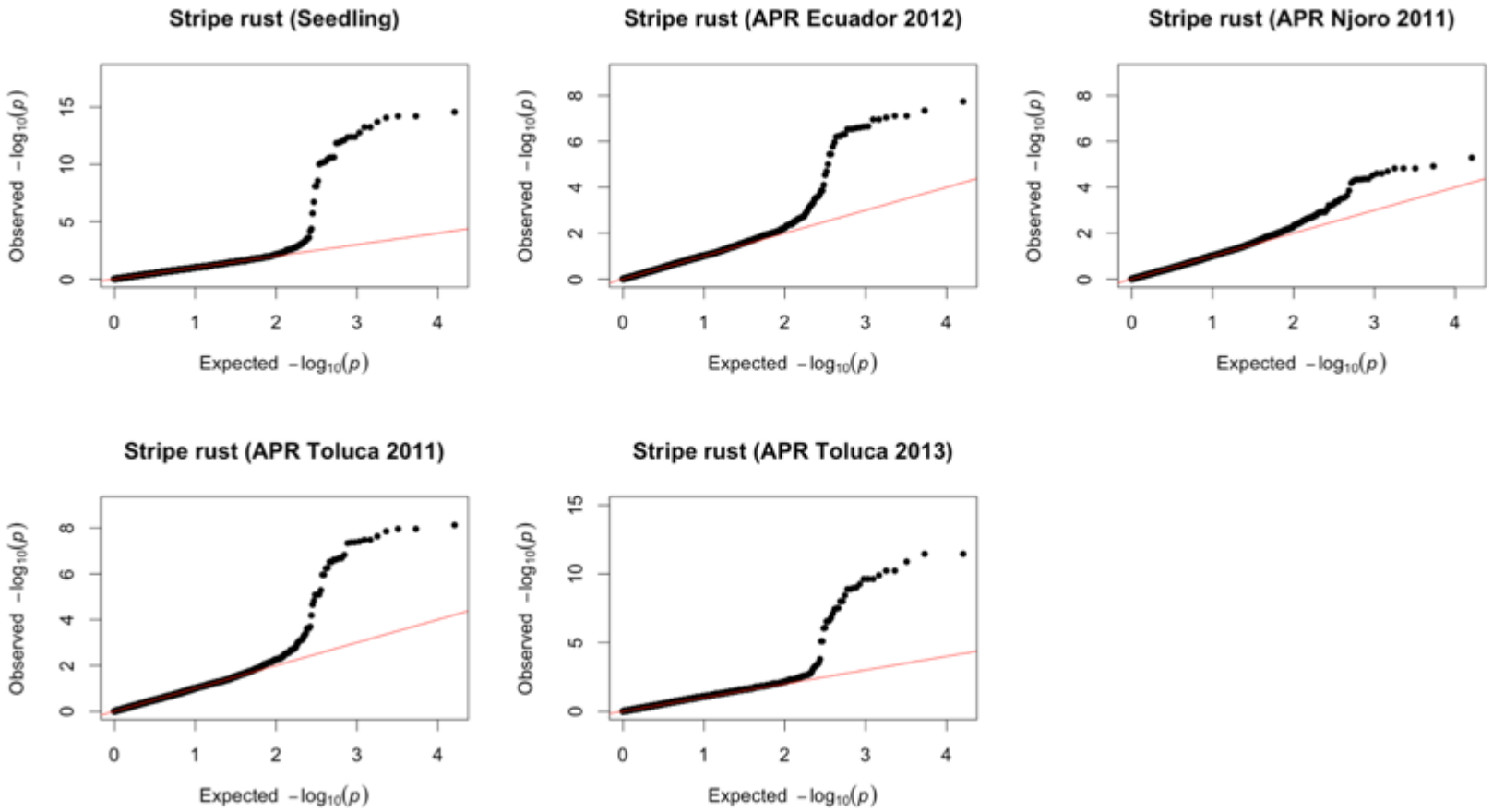

Figure S3.2: Quantile-quantile plot of the p-values comparing the p-value distribution to a uniform null distribution for seedling and adult plant resistance to stripe rust in the 46th IBWSN.